

Promotion of Open Science in Requirements Engineering – Leveraging the Open Research Knowledge Graph for FAIR Scientific Information

Oliver Karras, Alessio Ferrari, Davide Fucci, and Davide Dell'Anna

oliver.karras@tib.eu, alessio.ferrari@isti.cnr.it, davide.fucci@bth.se, d.dellanna@uu.nl

32nd IEEE International Requirements Engineering 2024 Conference – Exploring New Horizons:
Expanding the Frontiers of Requirements Engineering

June 24th, 2024, Reykjavik, Iceland

Welcome!



Oliver Karras



Alessio Ferrari



Davide Fucci



Davide Dell'Anna

Schedule and Table of Content

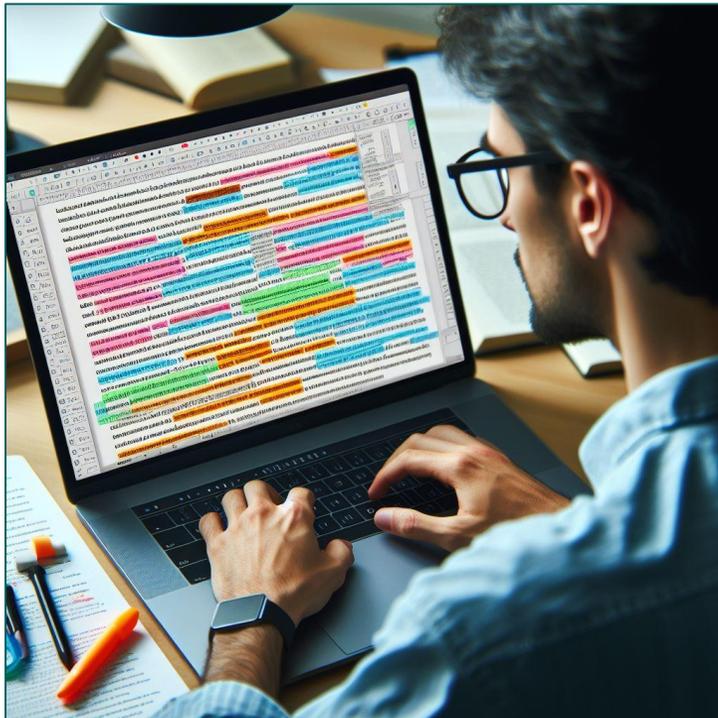
Session	Time	Table of Content	Style	Speaker
Theoretical	09:00 - 09:25	1. Welcome (5 min) 2. Introduction to open science in RE (10 min) 3. Introduction to the ORKG (10 min)	Presentation Presentation Presentation	All organizers Alessio Ferrari Oliver Karras
Practical	09:25 - 10:15	4. Create a FAIR-annotated publication for the ORKG (50 min) 4.1 Set up an Overleaf project for an exemplary publication 4.2 Use the LaTeX package SciKGT _E X to annotate the publication 4.3 Generate PDF with embedded FAIR scientific information 4.4 Optional: Upload the FAIR-annotated publication to the ORKG	Exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise	Oliver Karras All organizers All organizers All organizers All organizers
Break	10:15 - 10:45	Coffee break		
Practical	10:45 - 11:45	5. Use the ORKG based on a RE use case (60 min) 5.1 Add an exemplary publication to the ORKG 5.2 Describe the scientific information of the publication in the ORKG 5.3 Create an ORKG comparison of the publications added by participants 5.4 Publish the created ORKG comparison as a citable digital artifact 5.5 Optional: Create visualizations for the created ORKG comparison 5.6 Optional: Retrieve the information with the SPARQL endpoint	Exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise	Oliver Karras All organizers All organizers All organizers All organizers All organizers All organizers
Feedback	11:45 - 12:15	6. Reflection of the tutorial with the participants (25 min) 7. Farewell and closing (5 min)	Discussion Presentation	All organizers All organizers

Teaser: REFSQ'25 Open Science Competition

Challenge 1:

Annotate your REFSQ'25 paper with SciKGTeX.

The accepted paper, best annotated with SciKGTeX, will be awarded the **Best ORKG Annotation Award** (prize: 100€).



SciKGTeX

Challenge 2:

Enrich your RESFQ'25 paper with an ORKG comparison.

The accepted paper, enriched with the best ORKG comparison, will be awarded the **Best ORKG Comparison Award** (prize: 200€).



Open
Research
Knowledge
Graph

Divide and Conquer the EmpiRE:
A Community-Maintainable Knowledge Graph of
Empirical Research in Requirements Engineering

Oliver Karas*, Felix Wernlein[†], Jill Klünder[†] and Sören Auer*[†]
*TIB - Leibniz Information Centre for Science and Technology, Hannover, Germany
Email: {oliver.karas, soeren.auer}@tib.eu
[†]Leibniz University Hannover, Hannover, Germany
Email: felix.wernlein@stud.uni-hannover.de, jill.kluender@inf.uni-hannover.de, auer@13s.de

A Comparison of Scientific Publications on the State of Empirical Research in Requirements Engineering and Software Engineering ★

November 2023 • Oliver Karas • Felix Wernlein • Jill Ann-Christin Klünder • Sören Auer

This comparison provides an overview of scientific publications that have investigated primary studies in requirements engineering and software engineering to give a snapshot of the "current" state of empirical research in requirements engineering and software engineering. In particular, the comparison shows for each publication (1) which research fields and topics were investigated, (2) whether and where the extracted and analyzed data is available, and (3) which method was used to determine the state, including further details about the respective method.

DOI: <https://doi.org/10.48366/R650023>

Properties	Empirical research in requirements engineering: trends and opportunities - 2016	Empirical research methodologies and studies in Requirements Engineering: How far did we come? - 2014	A Survey on Empirical Requirements Engineering Research Practices - 2012	Evidence-Based Structuring and Evaluation of Empirical Research in Requirements Engineering: Fundamentals, Framework, Research Map - 2010	An Annotated Requirements Data - Empirical
research problem	empirical research in requirements engineering	empirical research in requirements engineering	empirical research in requirements engineering	empirical research in requirements engineering	empirical
research field investigated	Requirements Engineering	Requirements Engineering	Requirements Engineering	Requirements Engineering	Req.
topic investigated	bibliographic metadata context data collection	bibliographic metadata research topic theory	context data analysis data collection	context research method result	

your personal use. Not for redistribution. The definitive version of record was published in the proceedings of 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), <https://doi.org/10.1109/ESEM56168.2023.10304795>.

Introduction: Open Science and FAIR in Software Engineering and Requirements Engineering

Oliver Karras, Alessio Ferrari, Davide Fucci, and Davide Dell'Anna

oliver.karras@tib.eu, alessio.ferrari@isti.cnr.it, davide.fucci@bth.se, d.dellanna@uu.nl

32nd IEEE International Requirements Engineering 2024 Conference – Exploring New Horizons: Expanding the Frontiers of Requirements Engineering

June 24th, 2024, Reykjavik, Iceland

What is Open Science?

Open science is the movement to make scientific research (including publications, data, physical samples, and software) and its dissemination **accessible** to all levels of society, amateur or professional.

- Open methodology
- Open source
- Open data
- Open access
- Open peer review (identities, reports, community participation)
- Open educational resources



[1]

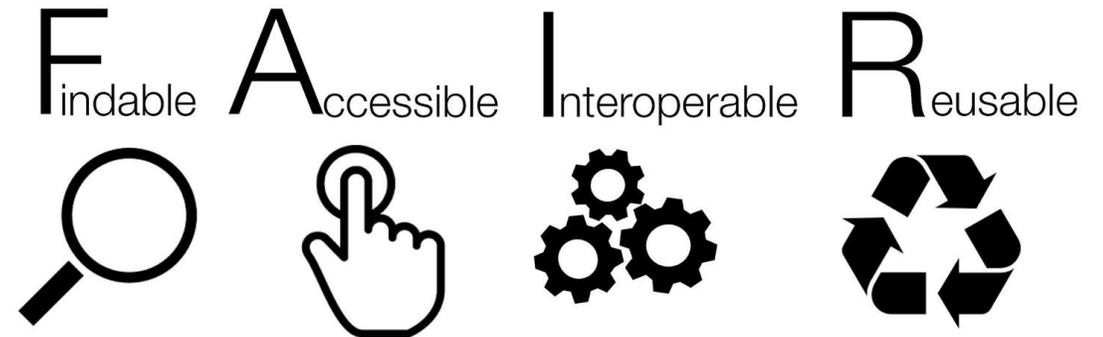
What are the FAIR Principles?

Scientists increasingly rely on computational support to deal with increasing amounts of data.

Focus on **Digital** Assets (Data & Knowledge)

FAIR Principles^[2]:

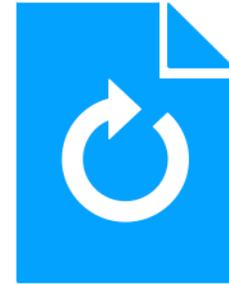
- **F**indability: Rich metadata, reachable by humans and machines, unique identifiers (DOI)
- **A**ccessibility: Open protocols for data access
- **I**nteroperability: Integration with other data
- **R**euse: Replicable experiments, understandable data (through metadata)



Why Do We Need Open Science and FAIR?

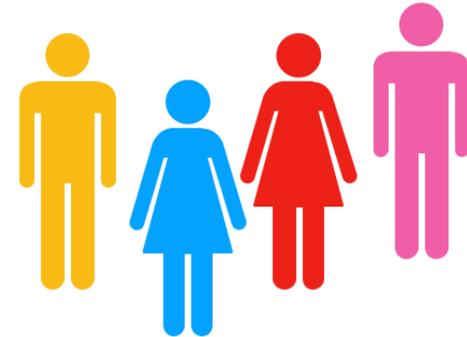
1. Transparency and Reproducibility:

- Open science helps verify the validity of research results
- Scientists can reproduce experiments, which is crucial for the **credibility**



2. Collaboration and Innovation:

- Collaboration among scientists, and **interdisciplinary** communication
- Accelerate discoveries and novel **ideas**



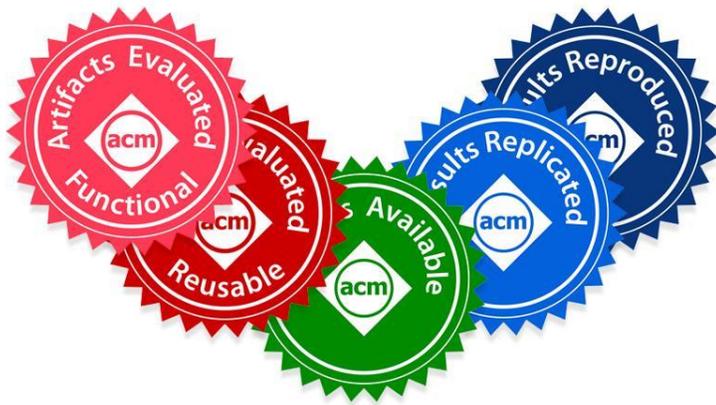
3. Equity and Inclusion:

- Democratize access to knowledge, and enrich viewpoints through **global access**
- Science accessible to the general public can increase scientific literacy, leading to an **informed society**



Current Initiatives in Software Engineering

Emphasis on Data, **Tools**, Research Results



The ACM badge system (2010)



Artifact Evaluation Tracks



Home > ACM Journals > ACM Transactions on Software Engineering and Methodology > CFP

Sections

Replicated Computational Results (RCR) Report

Journal Initiatives

Example Problems in RE: Replication and Tool Reconstruction

ID	Description	
Rec1	The reconstruction-relevant information and implementation details of the original approach can be ambiguous, imprecise, and incomplete.	Poor information
Rec2	If a tool was partially or fully developed and/or evaluated using proprietary data, then there is no guarantee that the reconstructed tool would be identical to the original one since the used data cannot be accessed for reconstruction purposes.	Proprietary data
Rec3	Communication with the original authors is not always useful since the actual information sources may not be available anymore.	Communication
Rec4	The continuous evolution of the NLP ecosystem entails that some libraries become outdated, unavailable, or not maintained anymore.	Updates and maintenance
Rec5	Tools are typically developed as prototypes and not maintained in the long term.	
Rec6	Tool reconstruction is not (yet) valued as a self-standing research contribution, and hence researchers are discouraged to replicate tools overtime.	Poor value

Example Problems in RE: Dataset Annotation

ID	Description
Ann1	Some RE-specific categorization tasks lack solid theories that can guide the annotation process.
Ann2	Besides annotation experience and theoretical knowledge, the lack of domain knowledge can limit the accuracy of the annotations.
Ann3	The annotation activity is time consuming due to factors such as language barriers, different individuals' background, and fatigue.
Ann4	The annotation protocol can evolve and thereby necessitate the re-annotation of the data which might, again, cause additional time and effort.
Ann5	Theoretical and practical training resources and opportunities are limited and not adequate for training novice annotators who are often trained during the annotation task by more experienced annotators.
Ann6	The lack of benchmarks entails that annotated datasets enabling comparison against the state-of-the-art are scarce.
Ann7	Available imbalanced datasets pose the challenge of both understanding the minority class and consequently the annotation of new examples thereof.
Ann8	Determining the right amount of context to be shared alongside the annotation raw data is essential and can significantly affect the annotation results.
Ann9	Motivating the annotators poses another challenge since an immediate observation of the impact of a given annotation task is not always possible.
Ann10	Annotators are often not experienced in managing the social aspects or resolving conflicts originating from power, authority, or other social relations.

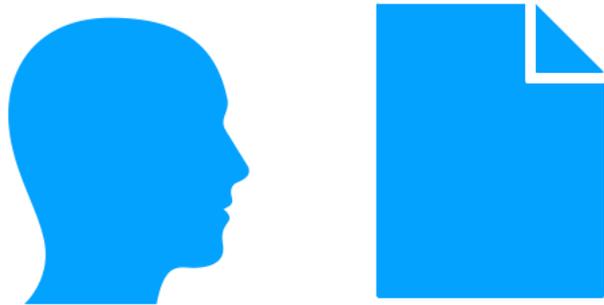
Domain Knowledge

Guidelines

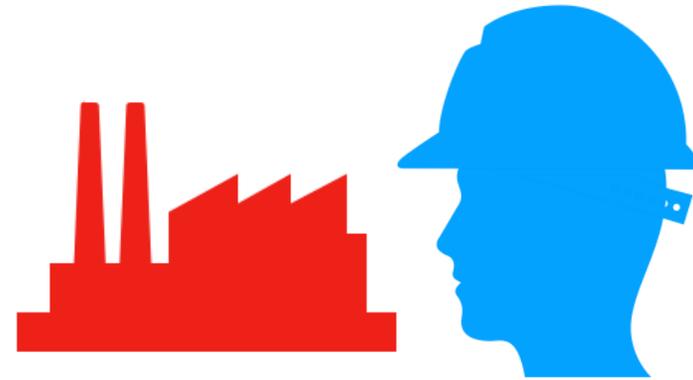
Data Quality

Lack of Benchmarks

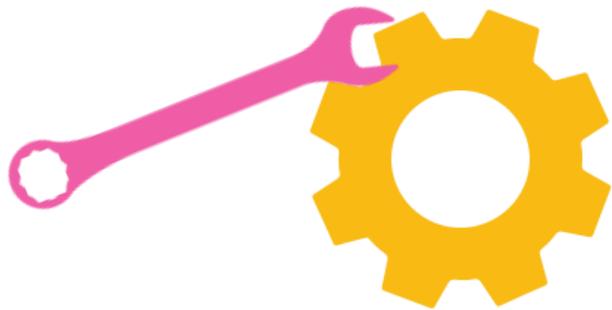
What do We Need?



Clear papers, Clear Communication



Involvement of companies and experts



Tool maintenance



Reward for open science and replication

Thanks to...



Sallam Abualhaija



Fatma Basak Aydemir



Fabiano Dalpiaz



Davide Dell'Anna



Xavier Franch



Davide Fucci

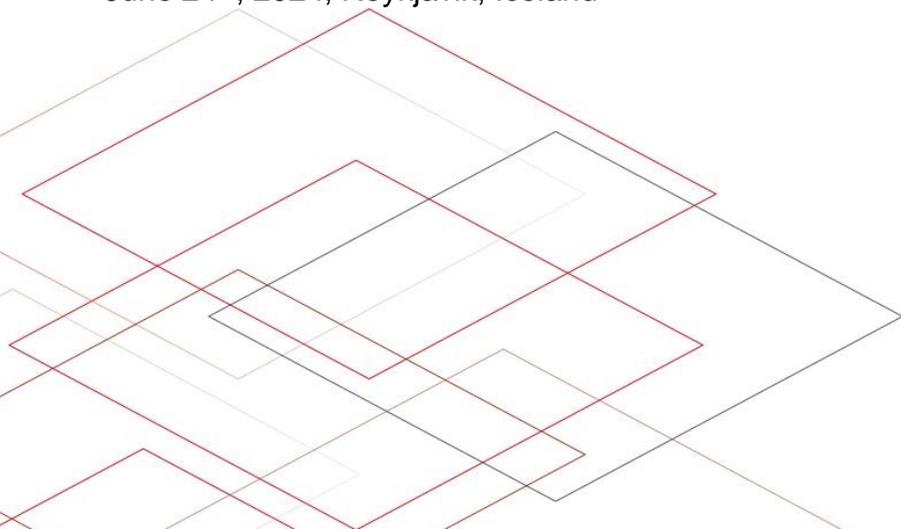
Introduction: Open Research Knowledge Graph

Oliver Karras, Alessio Ferrari, Davide Fucci, and Davide Dell'Anna

oliver.karras@tib.eu, alessio.ferrari@isti.cnr.it, davide.fucci@bth.se, d.dellanna@uu.nl

32nd IEEE International Requirements Engineering 2024 Conference – Exploring New Horizons: Expanding the Frontiers of Requirements Engineering

June 24th, 2024, Reykjavik, Iceland

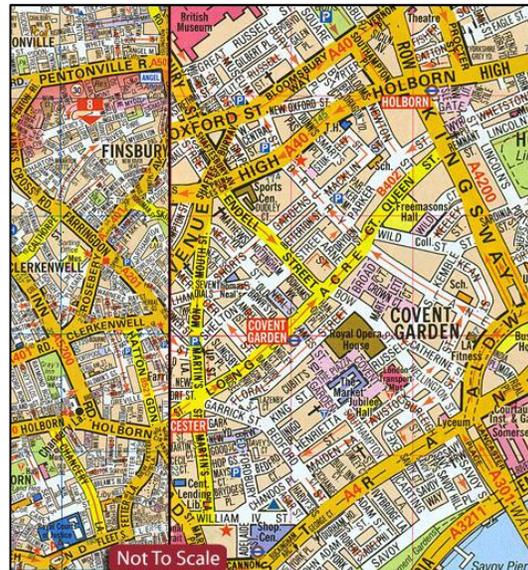


Once Upon a Time, we Communicated with Paper...

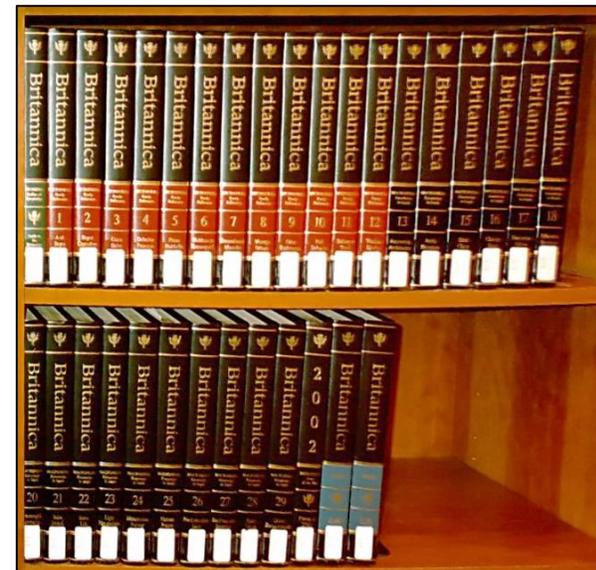
Who still remembers?



Mail order catalogs



Maps

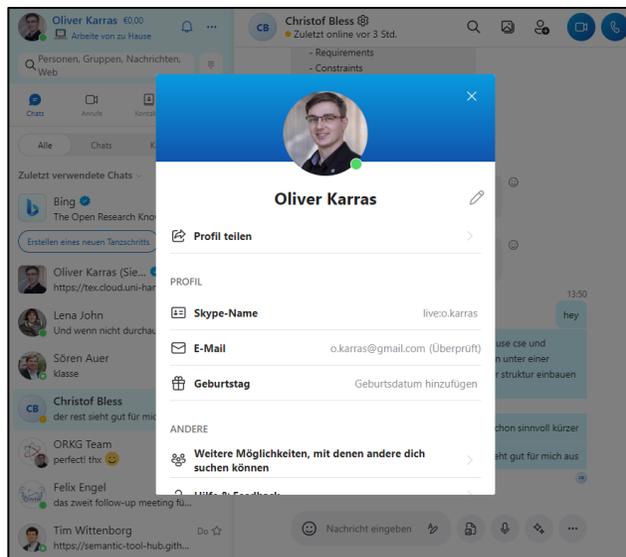
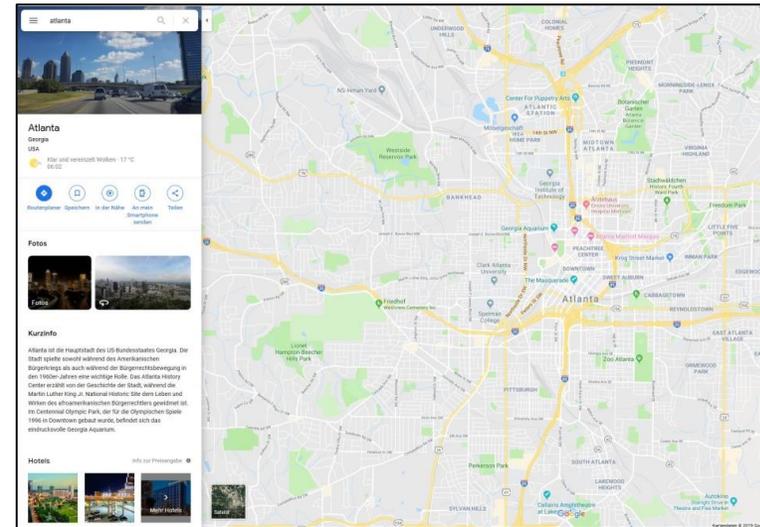
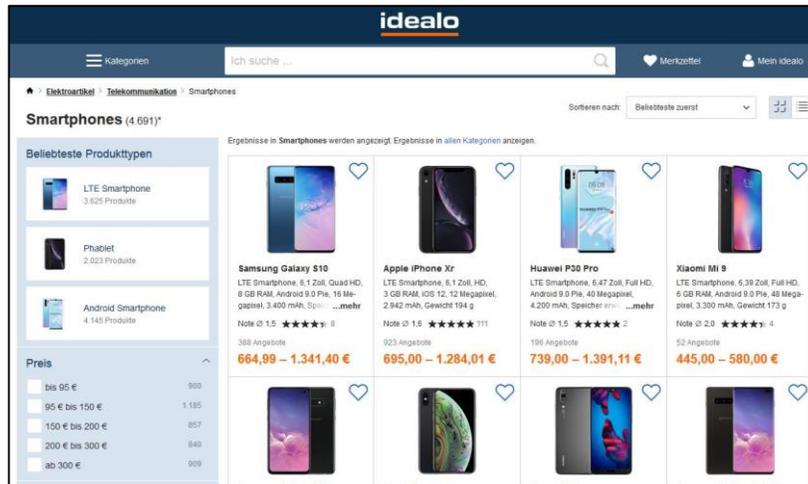


Encyclopedia



Phone books

... until Digital Transformation (Digitalization)!



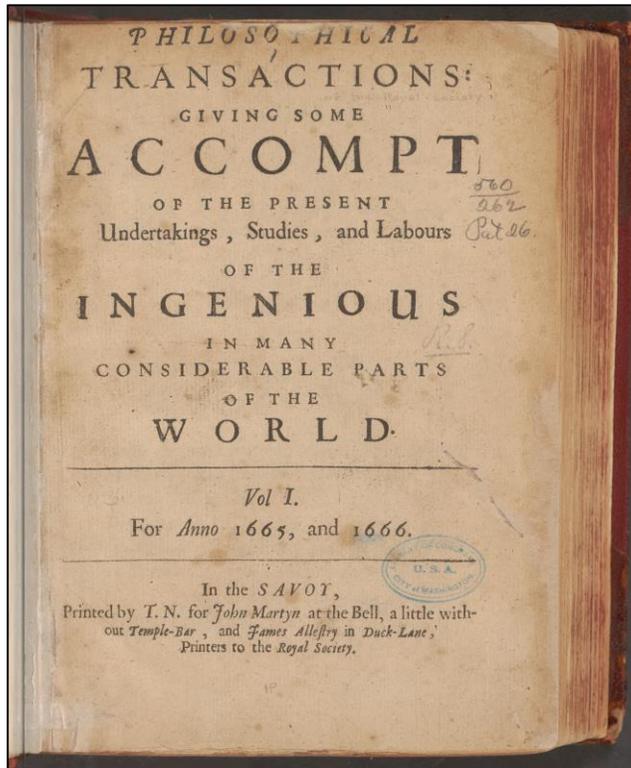
The World of Publishing & Communication has Profoundly Changed!

- **New means adapted to the new possibilities**, e.g., platforms
- **Completely new business models**
- **More focus on data, interlinking, services, and search**
- **Integration, crowdsourcing, and data curation** are important

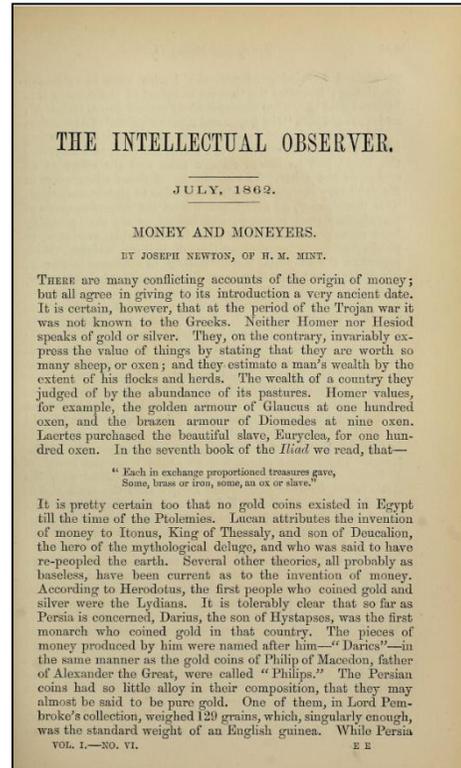
What has happened in academia in terms of scholarly publishing & communication?

Let's Take a Look

17th century



19th century



20th century

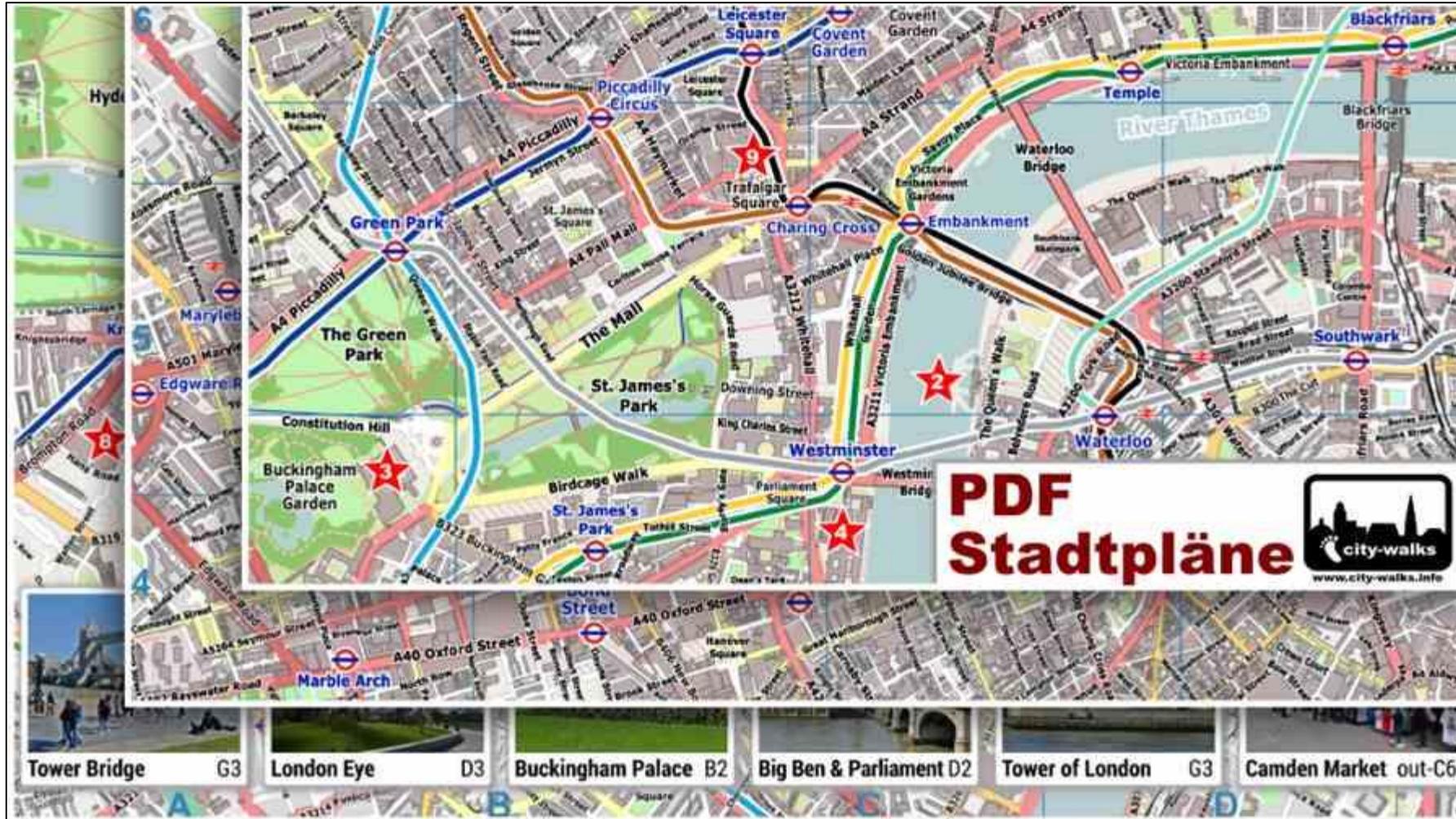


21st century



Scholarly publishing & communication has **not** changed (much)!

Let's Take a Look



21st century

AGDISTIS - Graph-Based Disambiguation of Named Entities Using Linked Data

Ricardo Usbeck^{1,2}, Axel-Cyrille Ngonga Ngomo¹, Michael Röder^{1,2}, Daniel Gerber¹, Sandro Athaide Coelho³, Sören Auer⁴, and Andreas Both²

¹ University of Leipzig, Germany
² R&D, Unister GmbH, Germany
³ Federal University of Juiz de Fora, Brazil
⁴ University of Bonn & Fraunhofer IAIS, Germany
(usbeck,ngonga)@informatik.uni-leipzig.de

Abstract. Over the last decades, several billion Web pages have been made available on the Web. The ongoing transition from the current Web of unstructured data to the Web of Data yet requires scalable and accurate approaches for the extraction of structured data in RDF (Resource Description Framework) from these websites. One of the key steps towards extracting RDF from text is the disambiguation of named entities. While several approaches aim to tackle this problem, they still achieve poor accuracy. We address this drawback by presenting AGDISTIS, a novel knowledge-base-agnostic approach for named entity disambiguation. Our approach combines the Hypertext-Induced Topic Search (HITS) algorithm with label expansion strategies and string similarity measures. Based on this combination, AGDISTIS can efficiently detect the correct URIs for a given set of named entities within an input text. We evaluate our approach on eight different datasets against state-of-the-art named entity disambiguation frameworks. Our results indicate that we outperform the state-of-the-art approach by up to 29% F-measure.

1 Introduction

The vision behind the Web of Data is to provide a new machine-readable layer to the Web where the content of Web pages is annotated with structured data (e.g., RDEa [1]). However, the Web in its current form is made up of at least 15 billion Web pages.¹ Most of these websites are unstructured in nature. Realizing the vision of a usable and up-to-date Web of Data thus requires scalable and accurate natural-language-processing approaches that allow extracting RDF from such unstructured data. Three tasks play a central role when extracting RDF from unstructured data: named entity recognition (NER), named entity disambiguation (NED), also known as entity linking [16], and relation extraction (RE). For the first sentence of Example 1, an accurate named entity recognition approach would return the strings Barack Obama and Washington,

¹ Data gathered from <http://www.worldwidewebsize.com/> on January 4th, 2014.

Let's Take a Look



Rethink How Scientific Knowledge is Communicated



*“The lightbulb was **not** invented by improving the candle.”*

Oren Harari

Digitalization is **more** than just Digitization!

Current and future scientific challenges can not be tackled with an outdated communication system.

**Digitalize Knowledge,
Not Documents!**

Example: Requirements Engineering and Empirical Research

☰
Google
Scholar

🔍

📄 Artikel

Ungefähr 3.970.000 Ergebnisse (0,16 Sek.)

Beliebige Zeit

Seit 2024

Seit 2023

Seit 2020

Zeitraum wählen...

Nach Relevanz
sortieren

Nach Datum sortieren

Beliebige Sprache

Seiten auf Deutsch

Alle Typen

Übersichtsarbeiten

Patente
einschließen

Zitate einschließen

[\[HTML\]](#) **Empirical research in requirements engineering: trends and opportunities** [\[HTML\]](#) [springer.com](#)

[T Ambreen](#), [N Ikram](#), [M Usman](#), [M Niazi](#) - [Requirements Engineering](#), 2018 - Springer

... trends and future **research** directions. To represent a state-of-the-art of **requirements engineering**, along with various trends and opportunities of **empirical RE research**, we conducted a ...

☆ Speichern [🔗 Zitieren](#) Zitiert von: 125 [Ähnliche Artikel](#) [Alle 7 Versionen](#)

[\[HTML\]](#) **Empirical research methodologies and studies in Requirements Engineering: How far did we come?** [\[HTML\]](#) [sciencedirect.com](#)

[M Daneva](#), [D Damian](#), [A Marchetto](#), [O Pastor](#) - [Journal of systems and ...](#), 2014 - Elsevier

... Software **Engineering** paradigm. We summarize prior **empirical research** in RE and introduce the contributors to this special issue on **empirical research** methodologies and studies in ...

☆ Speichern [🔗 Zitieren](#) Zitiert von: 67 [Ähnliche Artikel](#) [Alle 12 Versionen](#)

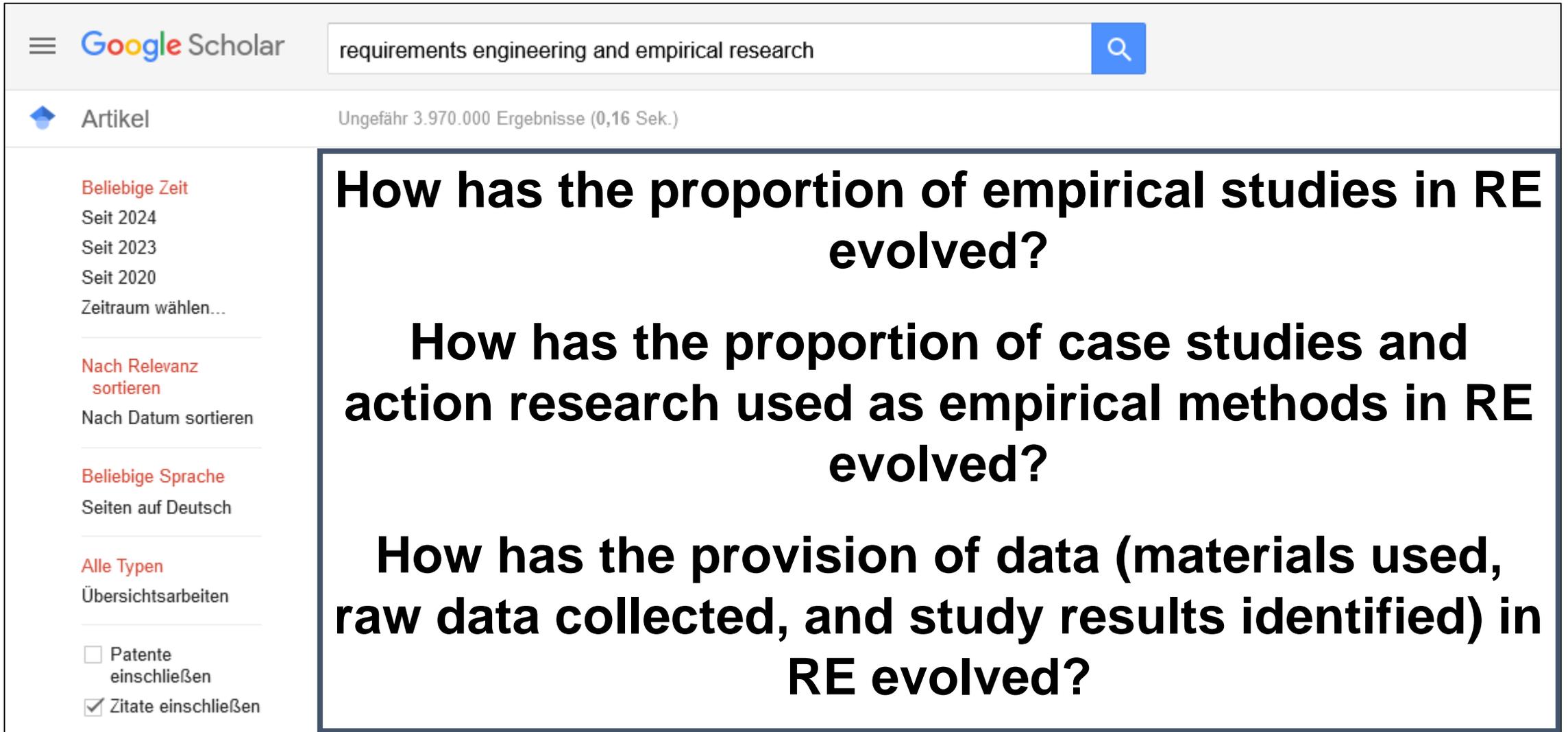
[\[PDF\]](#) **A systematic literature review of empirical research on quality requirements** [\[PDF\]](#) [springer.com](#)

[T Olsson](#), [S Sentilles](#), [E Papatheocharous](#) - [Requirements Engineering](#), 2022 - Springer

... We want to understand the **empirical research** on quality **requirements** topics as well as ... We conclude that more **research** is needed as **empirical research** on quality **requirements** is not ...

☆ Speichern [🔗 Zitieren](#) Zitiert von: 13 [Ähnliche Artikel](#) [Alle 9 Versionen](#)

Example: Requirements Engineering and Empirical Research



The image shows a screenshot of a Google Scholar search results page. The search query is "requirements engineering and empirical research". The results are filtered to "Artikel" (Articles) and show approximately 3,970,000 results in 0.16 seconds. The left sidebar contains filters for time period, relevance, language, and document type. The main content area is highlighted with a blue border and contains three research questions.

Google Scholar

requirements engineering and empirical research

Artikel

Ungefähr 3.970.000 Ergebnisse (0,16 Sek.)

Beliebige Zeit
Seit 2024
Seit 2023
Seit 2020
Zeitraum wählen...

Nach Relevanz sortieren
Nach Datum sortieren

Beliebige Sprache
Seiten auf Deutsch

Alle Typen
Übersichtsarbeiten

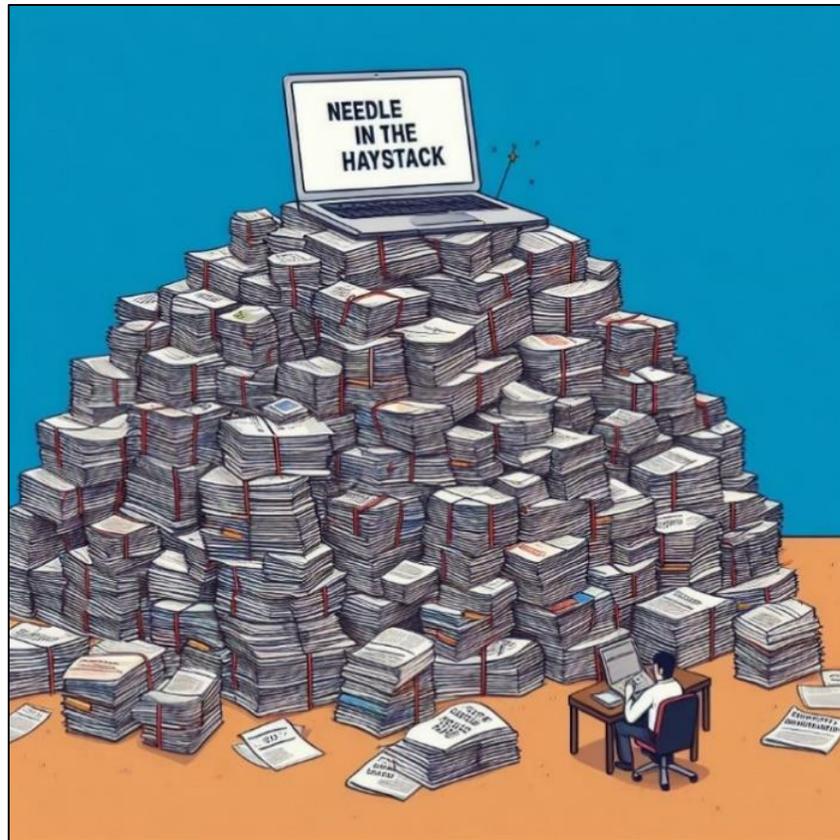
Patente einschließen
 Zitate einschließen

How has the proportion of empirical studies in RE evolved?

How has the proportion of case studies and action research used as empirical methods in RE evolved?

How has the provision of data (materials used, raw data collected, and study results identified) in RE evolved?

How do We Answer These Questions so far?

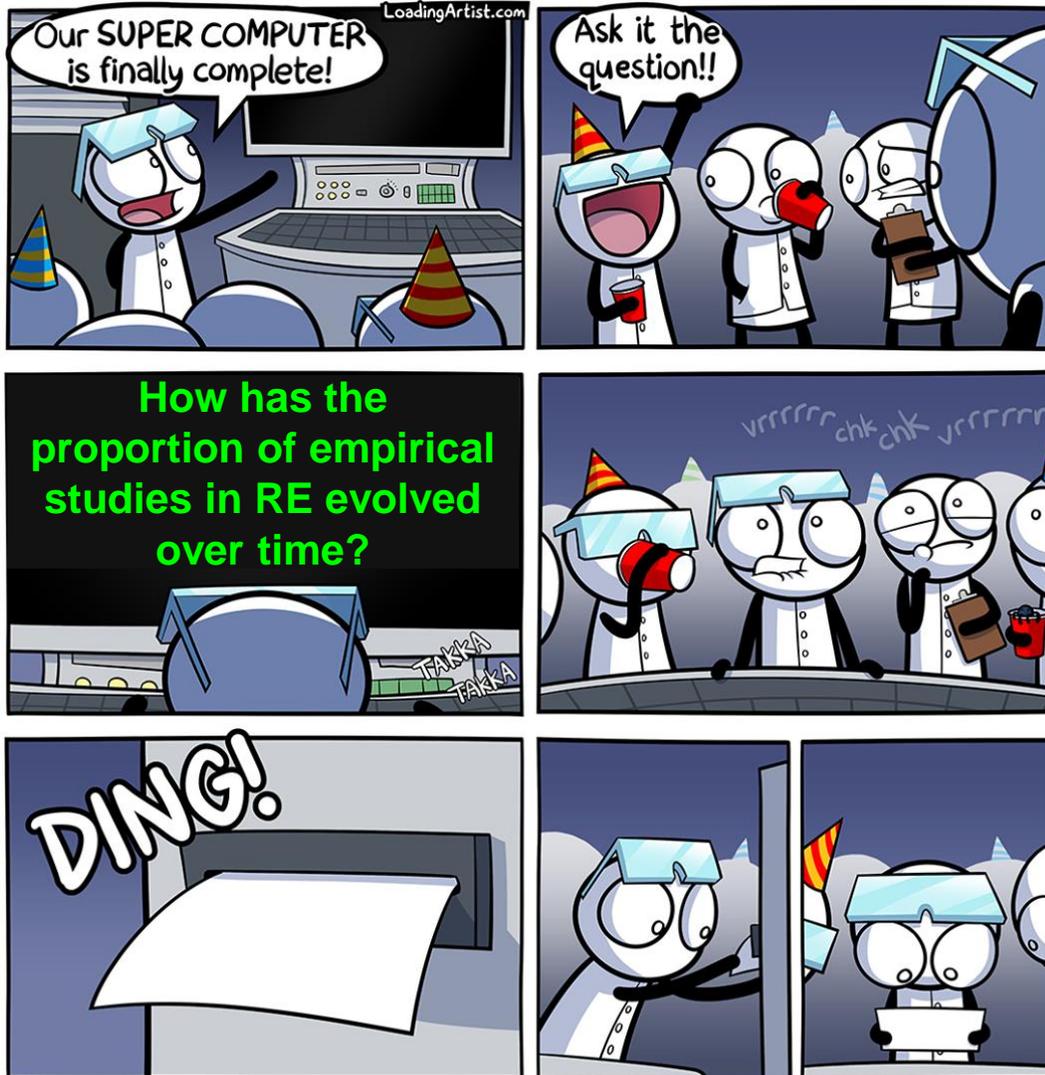


(S)LRs and SMSs for a **comprehensive, up-to-date, and long-term available** overview.



Over **7 million publications per year**^[4] with an **increasing** (exponential) tendency.

Wouldn't it be Great if we Could Ask the Computer?



How can we achieve this goal?

Open Research Knowledge Graph (ORKG)

The screenshot shows the ORKG website interface. At the top, there is a navigation bar with the ORKG logo, menu items (View, Tools, About), a dropdown for 'NFDI4DataScience', a search bar, and buttons for '+ Add new' and 'Sign in'. Below the navigation bar is a hero section with the text 'Scholarly Knowledge. FAIR.' and a sub-headline: 'The Open Research Knowledge Graph (ORKG) aims to describe research papers in a structured manner. With the ORKG, papers are easier to find and compare. Play video'. The main content area features a 'Browse by research field' section with five red buttons: 'Arts and Humanities' (432 papers - 33 comparisons), 'Engineering' (3391 papers - 345 comparisons), 'Life Sciences' (4142 papers - 195 comparisons), 'Physical Sciences & Mathematics' (15694 papers - 731 comparisons), and 'Social and Behavioral Sciences' (844 papers - 170 comparisons). Below this is a 'Comparisons' section with tabs for 'Comparisons', 'Papers', 'Visualizations', 'Reviews', and 'Lists'. The 'Comparisons' tab is active, showing a list of comparison items with details like 'Systematic Literature Review (SLR) Tools analysed based on General Features' (16 Contributions, 0 Visualizations, 04-03-2024) and 'RO estimates for infectious diseases' (6 Contributions, 0 Visualizations, 04-04-2024). On the right side, there are three widgets: 'Load Toots' (with a note about Mastodon), 'ORKG stories' (with a 'Find out more' link), and 'Join ORKG!' (with a 'Sign up' button). At the bottom right, there is a 'Knowledge base for science' section with a book icon and text: 'We are flooded with new publications in research every day and it is increasingly challenging to keep up'.

The ORKG revolutionizes how **scientific knowledge** is communicated and processed, making it **actionable** for machines and **researchers** to navigate, compare, and review **vast amounts of information** efficiently.

Research Knowledge Graphs

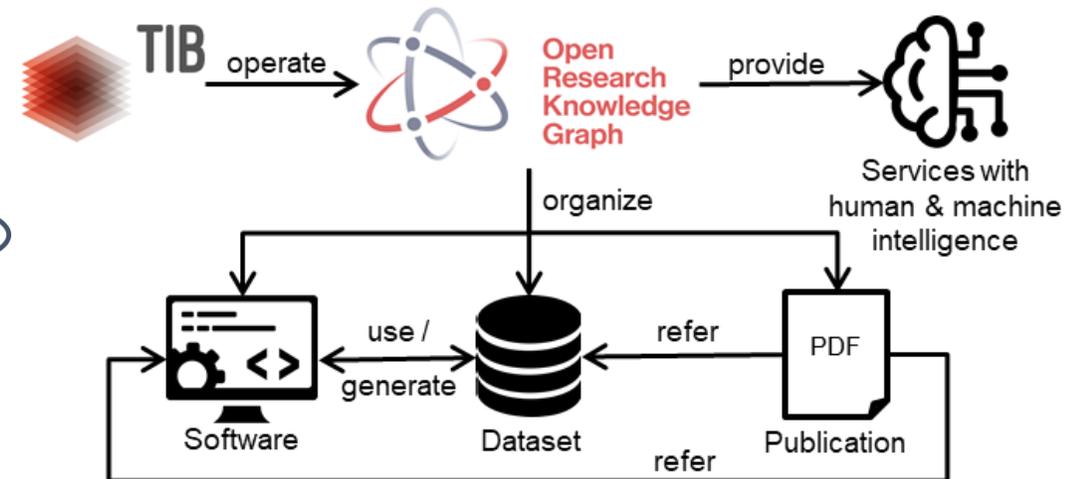
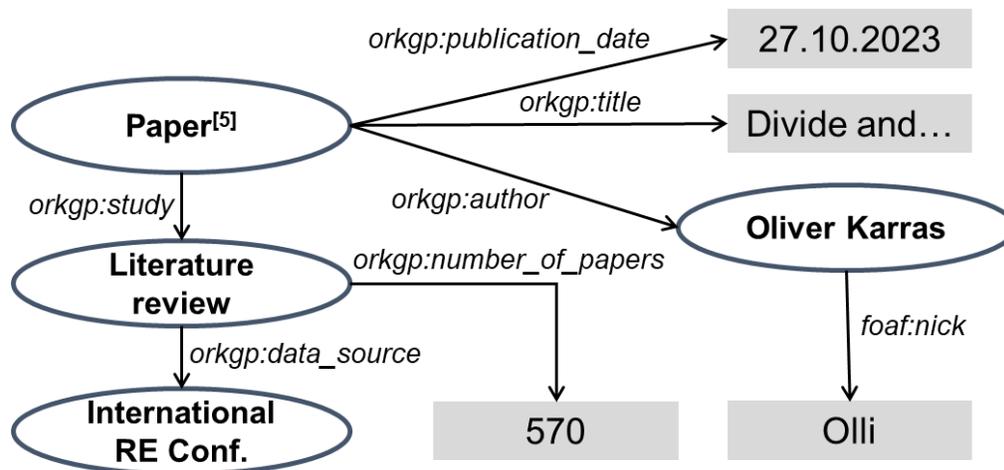
Research Knowledge Graphs (RKGs) are a **technology** for organizing scientific (meta-)data in a

- **Flexible, fine-grained, and semantic** representation
- That is **understandable** and **processable** by humans and machines

Research Knowledge Graphs

Research Knowledge Graphs (RKGs) are a **technology** for organizing scientific (meta-)data in a

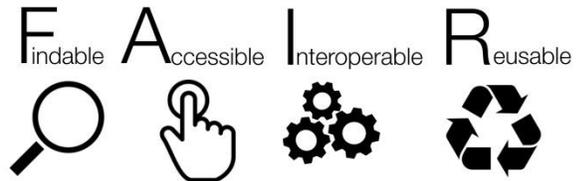
- **Flexible, fine-grained, and semantic** representation
- That is **understandable** and **processable** by humans and machines



The **ORKG** is a **ready-to-use** and **sustainably** operated **infrastructure** with **services** that uses a **cross-discipline RKG** for the **long-term** and **openly available organization** of scientific (meta-)data according to the **FAIR** data principles.

[5] Karras et al.: *Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering*. 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), DOI: [10.1109/ESEM56168.2023.10304795](https://doi.org/10.1109/ESEM56168.2023.10304795), 2023.

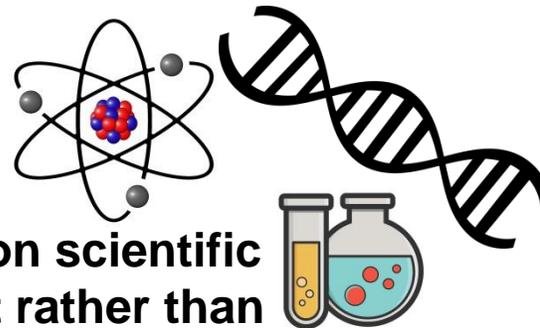
ORKG Objectives



Make research FAIR



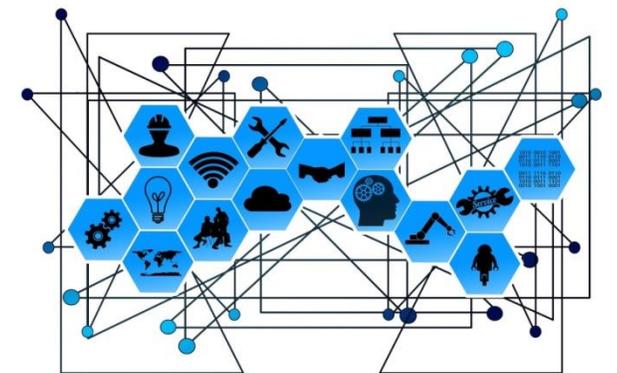
**Provide long-term
available and up-to-date
overviews**



**Focus on scientific
content rather than
pure metadata**



Foster collaboration



**Tackle interdisciplinary
challenges**

Structured Representation of Scientific Knowledge

Mining User Requirements from Application Store Reviews Using Frame Semantics

Nishant Jha and Anas Mahmoud^(✉)

The Division of Computer Science and Engineering, Louisiana State University,
Baton Rouge, LA 70803, USA
njha1@lsu.edu, mahmoud@csc.lsu.edu

Abstract. *Context and motivation:* Research on mining user reviews in mobile application (app) stores has noticeably advanced in the past few years. The majority of the proposed techniques rely on classifying the textual description of user reviews into different categories of technically informative user requirements and uninformative feedback. *Question/Problem:* Relying on the textual attributes of reviews often produces high dimensional models. This increases the complexity of the classifier and can lead to overfitting problems. *Principal ideas/results:* We propose a novel semantic approach for app review classification. The proposed approach is based on the notion of semantic role labeling, or characterizing the lexical meaning of text in terms of semantic frames. Semantic frames help to generalize from text (individual words) to more abstract scenarios (contexts). This reduces the dimensionality of the data and enhances the predictive capabilities of the classifier. Three datasets of user reviews are used to conduct our experimental analysis. Results show that semantic frames can be used to generate lower dimensional and more accurate models in comparison to text classification methods. *Contribution:* A novel semantic approach for extracting user requirements from app reviews. The proposed approach enables a more efficient classification process and reduces the chance of overfitting.

Keywords: Requirements elicitation · Application stores · Classification

1 Introduction

Mobile application markets, or app stores (e.g., Google Play and Apple App Store), represent a unique model of service-oriented business. Such platforms have created an unprecedented opportunity for app developers to directly monitor the opinions of a large population of end-users of their software [25]. Through app stores feedback services, app users can directly share their experience in the form of textual reviews and meta-data (e.g., star ratings). Analyzing large datasets of app store reviews has revealed that they contain substantial amounts of up-to-date technical information. Such information can be leveraged by app developers to help them maintain and sustain their apps in a highly-competitive

© Springer International Publishing AG 2017
P. Grünbacher and A. Perini (Eds.): REFSQ 2017, LNCS 10153, pp. 273–287, 2017.
DOI: 10.1007/978-3-319-54045-0_20



Mining User Requirements from Application Store Reviews Using Frame Semantics

📅 2017 Software Engineering 👤 Nishant Jha 👤 Anas Mahmoud

Published in: *Requirements Engineering: Foundation for Software Quality* DOI: 10.1007/978-3-319-54045-0_20

User Feedback Classification (Template) Add to comparison

Contribution data

← Back User *has experiment* → Machine Learning Experiment 🔗

Classification category	Feature request
Has dataset	seel.cse.lsu.edu/data/refsq17.zip
Has feature	Bag of words
Has method	Support vector machines
Has result	P = 0.45, R = 0.68, F1 = 0.54, etc.

Provenance **Timeline**

Added on
12 May 2021

Added by
 Oliver Karras

Contributors
Kheir Eddine
Eduard C. Groen
Oliver Karras

 Assign to observatory

Using FAIR Scientific Knowledge

How has the provision of data (materials used, raw data collected, and study results identified) in RE evolved?

Mining User Requirements from Application Store Reviews Using Frame Semantics

Nishant Jha and Anas Mahmoud^(✉)

The Division of Computer Science and Engineering, Louisiana State University,
Baton Rouge, LA 70803, USA
njha1@lsu.edu, mahmoud@csc.lsu.edu

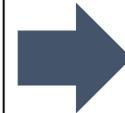
Abstract. *Context and motivation:* Research on mining user reviews in mobile application (app) stores has noticeably advanced in the past few years. The majority of the proposed techniques rely on classifying the textual description of user reviews into different categories of technically informative user requirements and uninformative feedback. *Question/Problem:* Relying on the textual attributes of reviews often produces high dimensional models. This increases the complexity of the classifier and can lead to overfitting problems. *Principal ideas/results:* We propose a novel semantic approach for app review classification. The proposed approach is based on the notion of semantic role labeling, or characterizing the lexical meaning of text in terms of semantic frames. Semantic frames help to generalize from text (individual words) to more abstract scenarios (contexts). This reduces the dimensionality of the data and enhances the predictive capabilities of the classifier. Three datasets of user reviews are used to conduct our experimental analysis. Results show that semantic frames can be used to generate lower dimensional and more accurate models in comparison to text classification methods. *Contribution:* A novel semantic approach for extracting user requirements from app reviews. The proposed approach enables a more efficient classification process and reduces the chance of overfitting.

Keywords: Requirements elicitation · Application stores · Classification

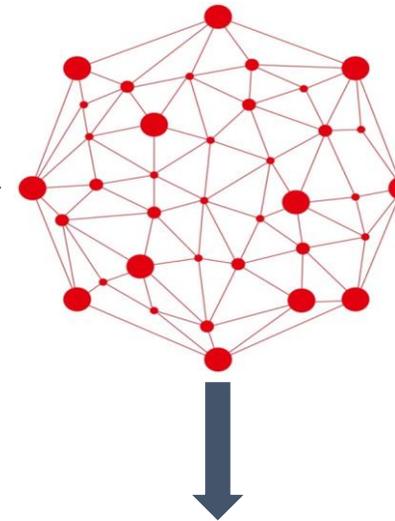
1 Introduction

Mobile application markets, or app stores (e.g., Google Play and Apple App Store), represent a unique model of service-oriented business. Such platforms have created an unprecedented opportunity for app developers to directly monitor the opinions of a large population of end-users of their software [25]. Through app stores feedback services, app users can directly share their experience in the form of textual reviews and meta-data (e.g., star ratings). Analyzing large datasets of app store reviews has revealed that they contain substantial amounts of up-to-date technical information. Such information can be leveraged by app developers to help them maintain and sustain their apps in a highly-competitive

© Springer International Publishing AG 2017
P. Grünbacher and A. Perini (Eds.): REFSQ 2017, LNCS 10153, pp. 273–287, 2017.
DOI: 10.1007/978-3-319-54045-0_20

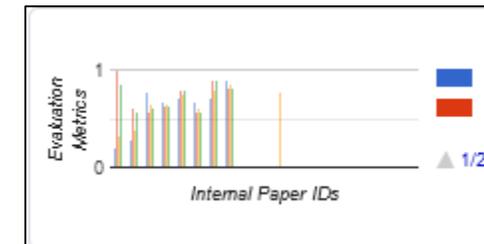


Natural
language
question



Answer

Properties	Software Feature Request Detection in Issue Tracking Systems <small>2016 - User Feedback Classification</small>	Mining User Requirements from Application Store Reviews Using Frame Semantics <small>2017 - User Feedback Classification</small>
Has dataset ▼	https://zenodo.org/record/56907#.YKT_NudCRPY	https://mast.informatik.uni-hamburg.de/wp-content/uploads/2014/03/REJ_data.zip https://sites.google.com



ORKG Comparisons

Acknowledgement of creators

DOI

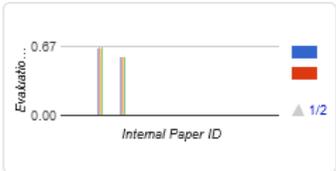
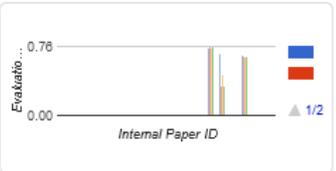
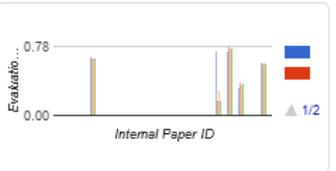
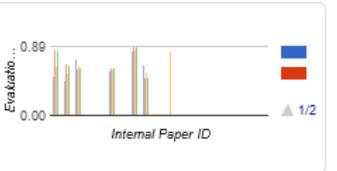
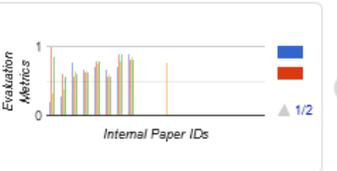
Visualizations

Interactive filtering

Overview of Approaches that Classify User Feedback as Feature Request ☆ 👁

📅 June 2021
👤 Oliver Karras
👤 Eduard C. Groen

This overview shows the classification results of approaches that use the machine learning algorithms Naïve Bayes, Support Vector Machines, and Decision Trees C4.5 in combination with the machine learning features Bag of Words or Term Frequency - Inverse Document Frequency to classify user feedback as feature request.

Properties	Software Feature Request Detection in Issue Tracking Systems <i>User Feedback Classification - 2016</i>	Mining User Requirements from Application Store Reviews Using Frame Semantics <i>User Feedback Classification - 2017</i>	Mining Twitter Feeds for Software User Requirements <i>User Feedback Classification - 2017</i>	Automatic Classification of Non-Functional Requirements from Augmented App User Reviews <i>User Feedback Classification - 2017</i>	Bug reports simply classified <i>User Feedback Classification - 2017</i>
has dataset	https://zenodo.org/record/56907#.YKT_NudCRPY	https://mast.informatik.uni-hamburg.de/wp-content/uploads/2014/03/REJ_data.zip https://sites.google.com/site/appsuserreviews/ seel.cse.lsu.edu/data/refsq17.zip	seel.cse.lsu.edu/data/re17.zip	Not available	https://seel.cse.lsu.edu/data/re17.zip

ORKG Comparisons are Citable

Comparison of Studies on Germany's Energy Supply in 2050 ★ 🔍

📅 November 2021
👤 Felix Kullmann
👤 Jan Göpfert
👤 Oliver Karras
👤 Patrick Kuckertz
👤 Sören Auer
👤 Markus Stocker

👤 Peter Markewitz
👤 Leander Kotzur
👤 Detlef Stolten



Share
🌐
🐦
🌐
🔗

This comparison compares the results of various studies analyzing a future low-carbon energy system for Germany. The focus of this study comparison is electricity generation. In the future, however, other essential characteristics of the respective energy system designs in the individual studies will be listed. Installed capacity is given in GW and electricity generation is given in TWh.

DOI: <https://doi.org/10.21203/rs.3.rs-1234567/v1>



Oliver Karras

Bearbeiten
Löschen

Comparison of studies on Germany's energy supply in 2050

Autoren Felix Kullmann, Peter Markewitz, Detlef Stolten, Oliver Karras, Patrick Kuckertz, Leander Kotzur, Jan-Maris Göpfert, Sören Auer, Markus Stocker

Publikationsdatum 2021

Ausgabe FZJ-2022-00782

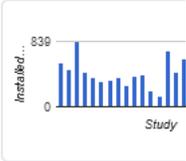
Verlag Technoökonomische Systemanalyse

Beschreibung This comparison compiles the results from various studies analyzing a future low-carbon energy system for Germany. The focus of this study comparison is electricity generation. In the future, however, other essential characteristics of the respective energy system designs in the individual studies will be listed. Installed capacity is given in GW and electricity generation is given in TWh.

Zitate insgesamt Zitiert von: 2

2023 2024

Google Scholar-Artikel [Comparison of studies on Germany's energy supply in 2050](#)
F Kullmann, P Markewitz, D Stolten, O Karras... - 2021
[Zitiert von: 2](#) [Ähnliche Artikel](#)



Den Weg zu einem treibhausgasneutralen Deutschland ressourcenschonend gestalten
Contribution 1 - 2019

- [all sources](#)
- [bioenergy](#)
- [geothermics](#)
- [hydropower](#)
- [import](#)
- [net import](#)

Properties

has energy sources

Klimaneutralität
Contributions

net import

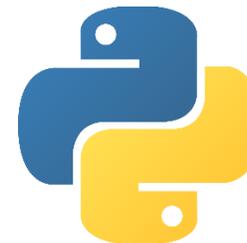
So far so Good, but...

...what can we do with machine-actionable scientific knowledge?

Anything we want!

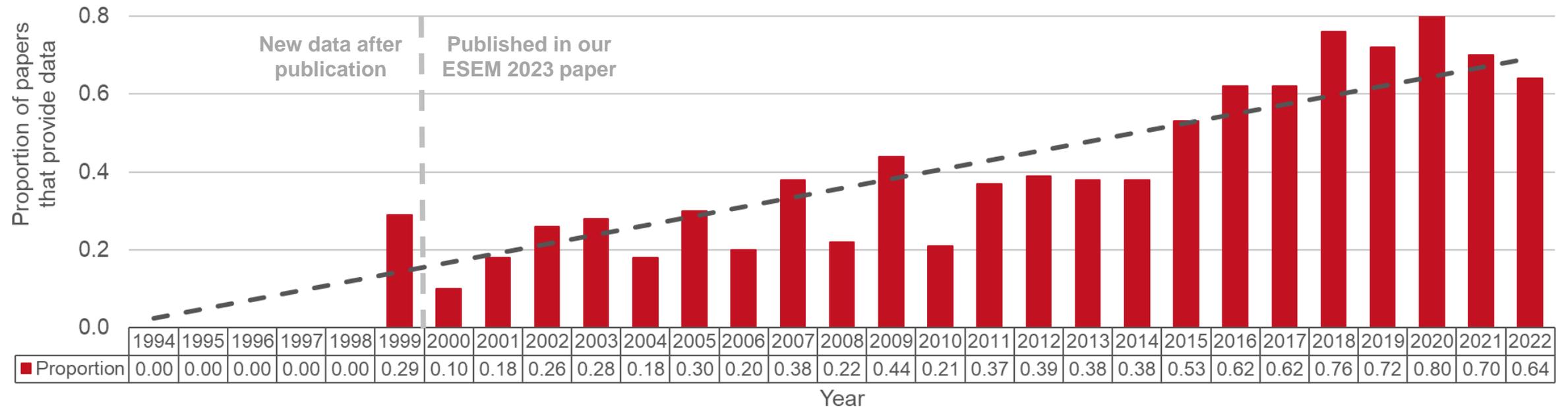
1. Papers, comparisons, and visualizations in the ORKG are **openly available for (re-)use and extension to anyone.**
2. The ORKG provides several **interfaces** for processing the data, e.g., to develop novel search, retrieval, mining, and assistance applications.

{ REST : API }



(Re-)use and Extension of Scientific Knowledge

How has the provision of data (materials used, raw data collected, and study results identified) in RE evolved? [5, 6, 7]
(Based on 680 papers in the ORKG)



[5] Karras et al.: *Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering*. 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), DOI: [10.1109/ESEM56168.2023.10304795](https://doi.org/10.1109/ESEM56168.2023.10304795), 2023.

[6] Project on GitHub: <https://github.com/okarras/EmpiRE-Analysis>

[7] Interactive Jupyter notebook: <https://mybinder.org/v2/gh/okarras/EmpiRE-Analysis/HEAD?labpath=%2Fempire-analysis.ipynb>

ORKG Observatory

Empirical Software Engineering | Observatory Edit

This observatory works to provide a community-maintainable knowledge graph of empirical research in software engineering. Our goal is to continuously acquire and curate comprehensive knowledge about empirical research applied in scientific publications in the research field of software engineering and its subfields, such as requirements engineering. In this way, we want to provide a comprehensive, up-to-date, and long-term overview of the state-of-the-art on empirical research in software engineering.

We are currently working on a knowledge graph of empirical research in requirements engineering.

For this purpose, we have developed a corresponding ORKG template (cf. <https://orkg.org/template/R186491>).

Contact: oliver.karras@tib.eu

Research problems

- empirical research in software engineering
- empirical research in requirements engineering

Organizations



Members

 **Oliver Karras**
TIB - Leibniz Information Centre for Science and Technology

 **Jil Kluender**
Leibniz University Hannover

Content | 688 items

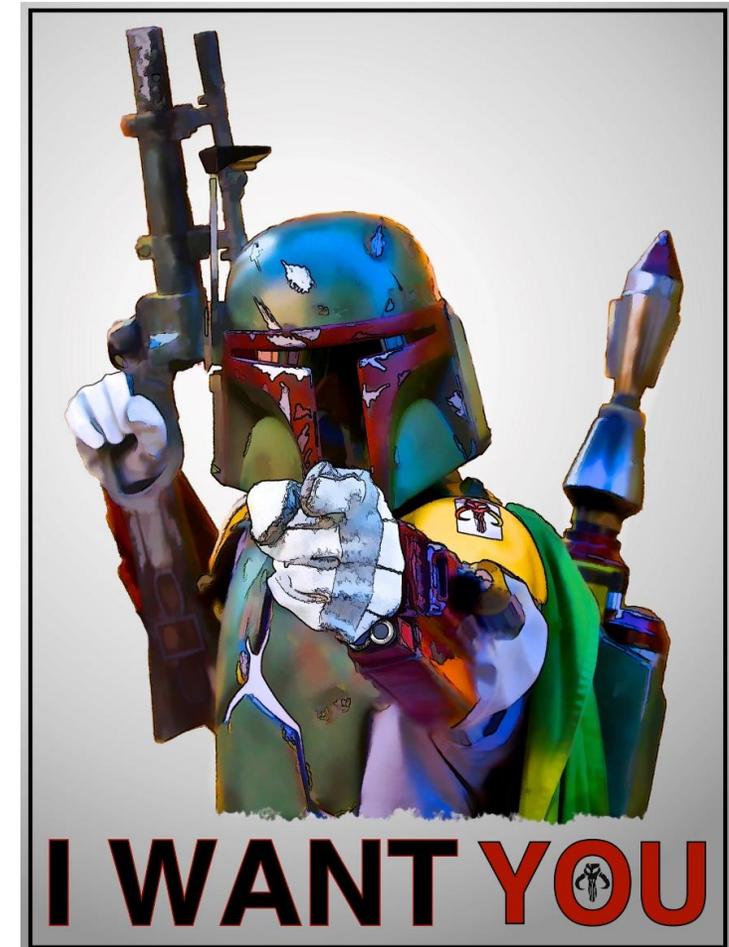
Show: Paper Comparison Visualization Top recent

- **Open** groups maintain topics in the ORKG
- **Central** access point for the community to all curated contents

Who organizes this scientific knowledge in the ORKG?

Who Organizes this Scientific Knowledge in the ORKG?

The heart of the ORKG is its **CROWD**:
Researchers from any discipline!



How can We Contribute to the ORKG?

1. While writing a paper:

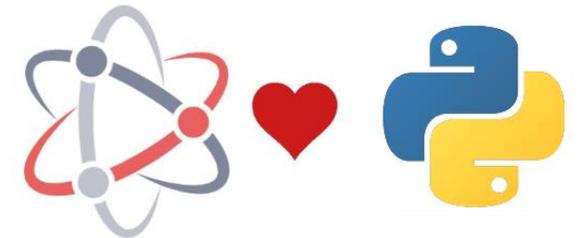
- SciKGTex: LaTeX package for FAIR annotations in a paper (embedded in the PDF) that can be imported into ORKG
- <https://github.com/Christof93/SciKGTex>



SciKGTex

2. While developing an analysis:

- Python package (& R package in development) for FAIR annotations in analysis scripts that can be imported into the ORKG
- <https://pypi.org/project/orkg/>



3. At any time:

- Manually using the ORKG Frontend to describe papers
- (Semi-)automatically using the ORKG REST API



Open
Research
Knowledge
Graph

{ REST : API }

What We Will Learn to Contribute to the ORKG

1. While writing a paper:

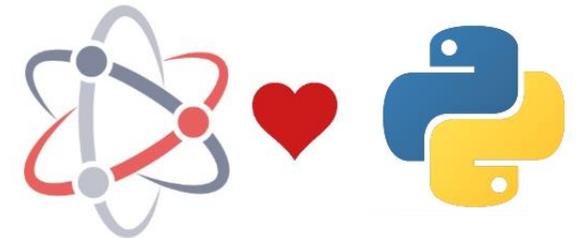
- **SciKGTex**: LaTeX package for FAIR annotations in a paper (embedded in the PDF) that can be imported into ORKG
- <https://github.com/Christof93/SciKGTex>



SciKGTex

2. While developing an analysis:

- Python package (& R package in development) for FAIR annotations in analysis scripts that can be imported into the ORKG
- <https://pypi.org/project/orkg/>



3. At any time:

- **Manually** using the ORKG Frontend to describe papers
- (Semi-)automatically using the ORKG REST API



Open
Research
Knowledge
Graph

{ REST : API }

Conclusion



Mining User Requirements from Application Store Reviews Using Frame Semantics

Nishant Jha and Anas Mahmoud^(*)
The Division of Computer Science and Engineering, Louisiana State University,
Baton Rouge, LA 70803, USA
njha@lsu.edu, mahmoud@ce.lsu.edu

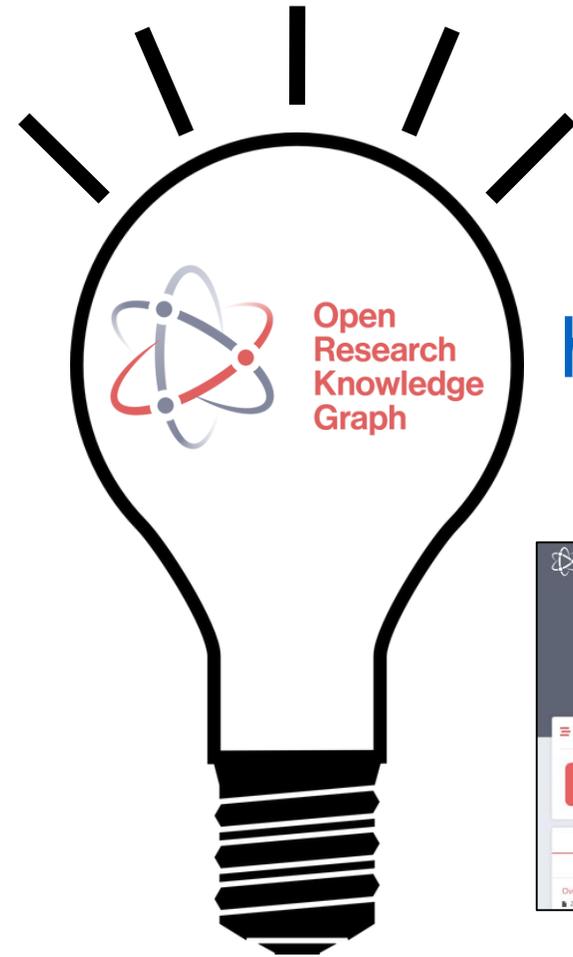
Abstract. Context and motivation: Research on mining user reviews in mobile application (app) stores has noticeably advanced in the past few years. The majority of the proposed techniques rely on classifying the textual description of user reviews into different categories of technically informative user requirements and uninformative feedback. Question/Problem: Relying on the textual attributes of reviews often produces high dimensional models. This increases the complexity of the classifier and can lead to overfitting problems. Principal idea/solution: We propose a novel semantic approach for app review classification. The proposed approach is based on the notion of semantic role labeling, or characterizing the lexical meaning of text in terms of semantic frames. Semantic frames help to generalize from text (individual words) to more abstract semantic (contexts). This reduces the dimensionality of the data and enhances the predictive capabilities of the classifier. Three datasets of user reviews are used to conduct our experimental analysis. Results show that semantic frames can be used to generate lower dimensional and more accurate models in comparison to text classification methods. Contribution: A novel semantic approach for extracting user requirements from app reviews. The proposed approach enables a more efficient classification process and reduces the chance of overfitting.

Keywords: Requirements elicitation · Application stores · Classification

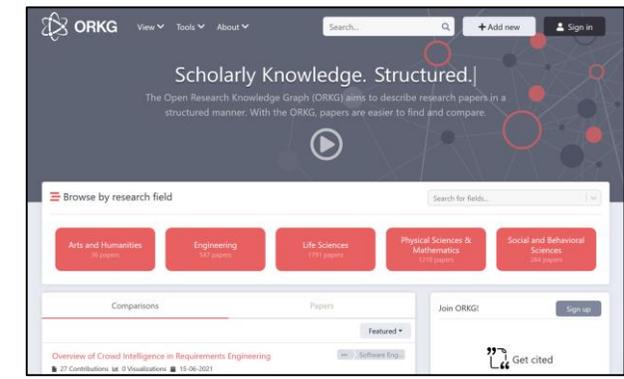
1 Introduction

Mobile application markets, or app stores (e.g., Google Play and Apple App Store), represent a unique model of service-oriented business. Such platforms have created an unprecedented opportunity for app developers to directly monitor the opinions of a large population of end-users of their software [2]. Through app stores feedback services, app users can directly share their experience in the form of textual reviews and meta-data (e.g., star ratings). Analyzing large datasets of app store reviews has revealed that they contain substantial amounts of up-to-date technical information. Such information can be leveraged by app developers to help them maintain and sustain their apps in a highly-competitive

© Springer International Publishing AG 2017
© Gölöskeci and A. Pász (Eds.), REEP 2017, LNCS 10153, pp. 273–287, 2017.
DOI: 10.1007/978-3-319-54054-2_20



<https://orkg.org>



Let's bring scholarly communication and open science in Requirements Engineering to the 21st century!

Further Reading



Open Access ORKG Book

- Celebrating the 5th anniversary of ORKG
- A practical guide for new and advanced users
 - ORKG's terms and concepts
 - ORKG's approach
 - ORKG's technology
 - ORKG's success stories

<https://cuvillier.de/de/shop/publications/9037-open-research-knowledge-graph>



SciKGT_eX: A LaTeX Package for FAIR-Annotated Publications

Oliver Karras, Alessio Ferrari, Davide Fucci, and Davide Dell'Anna

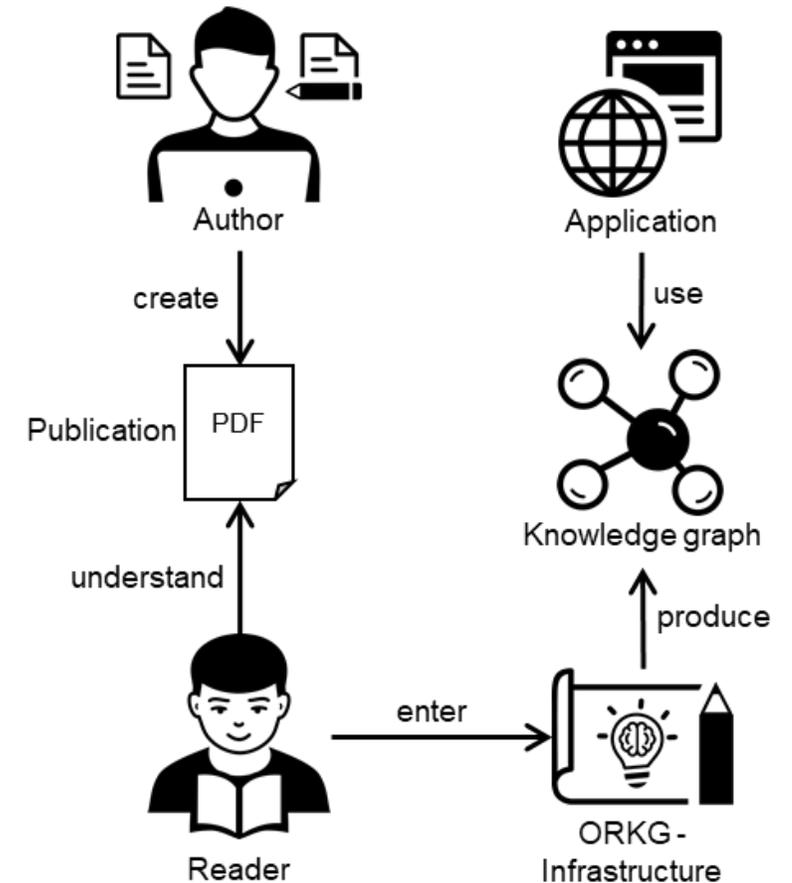
oliver.karras@tib.eu, alessio.ferrari@isti.cnr.it, davide.fucci@bth.se, d.dellanna@uu.nl

32nd IEEE International Requirements Engineering 2024 Conference – Exploring New Horizons: Expanding the Frontiers of Requirements Engineering

June 24th, 2024, Reykjavik, Iceland

Problem: Making Publications FAIR is a Downstream Task

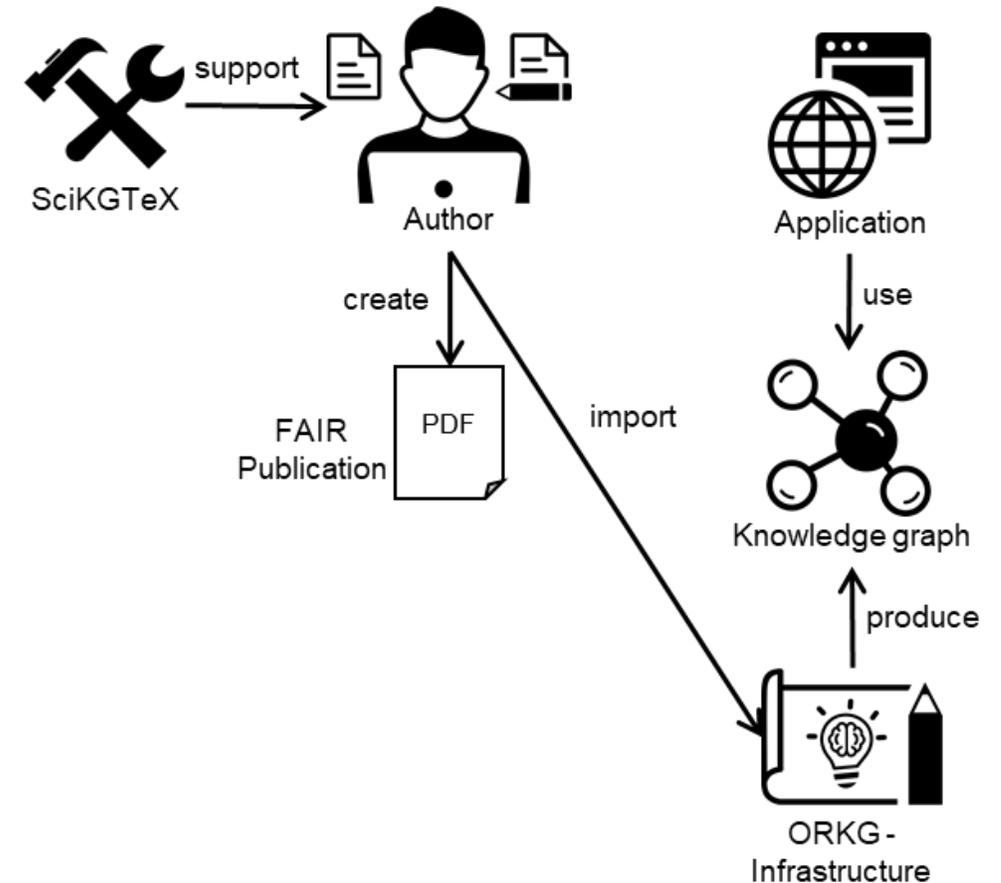
Problem: ORKG focuses on **published** articles, so making them **FAIR** is a **downstream task** and often **not done by the author**.



Solution: “FAIR-by-Design” Artifacts and SciKGTex^[8]

Solution: “FAIR-by-Design” Artifacts.
SciKGTex supports authors in making their publication **FAIR** at the time of **creation**.

- **Authors describe** their publication with FAIR information **only once** and **in parallel** at the time of **creation**
- **SciKGTex embeds** FAIR information into the **PDF metadata** as a knowledge graph
 - **Persistent** over PDF lifetime
 - **Available** for anyone
 - **Reusable**, e.g., import into ORKG



SciKGT_EX – Scientific Knowledge Graph TeX

- Predefined commands for annotation
 - 3 commands for metadata
 - Title
 - Author
 - Research field
 - 5 commands for content
 - Research problem
 - Objective
 - Method
 - Result
 - Conclusion
- Support for own custom annotations
 - REFSQ'24 and REFSQ'25 ask for
 - Code repository
 - Dataset

```
\usepackage{scikgtex}
```

```
\begin{document}
```

The role of `\researchproblem{antibiotic therapy}` is controversial. The purpose of this study was to `\objective{determine the effectiveness of high-dose amoxicillin/potassium clavulanate in the treatment of children}`.

This was a `\method{randomized, double-blind, placebo-controlled study}`.

`\result{Children receiving the antibiotic were more likely to be cured (50% vs 14%) than children receiving the placebo}`.

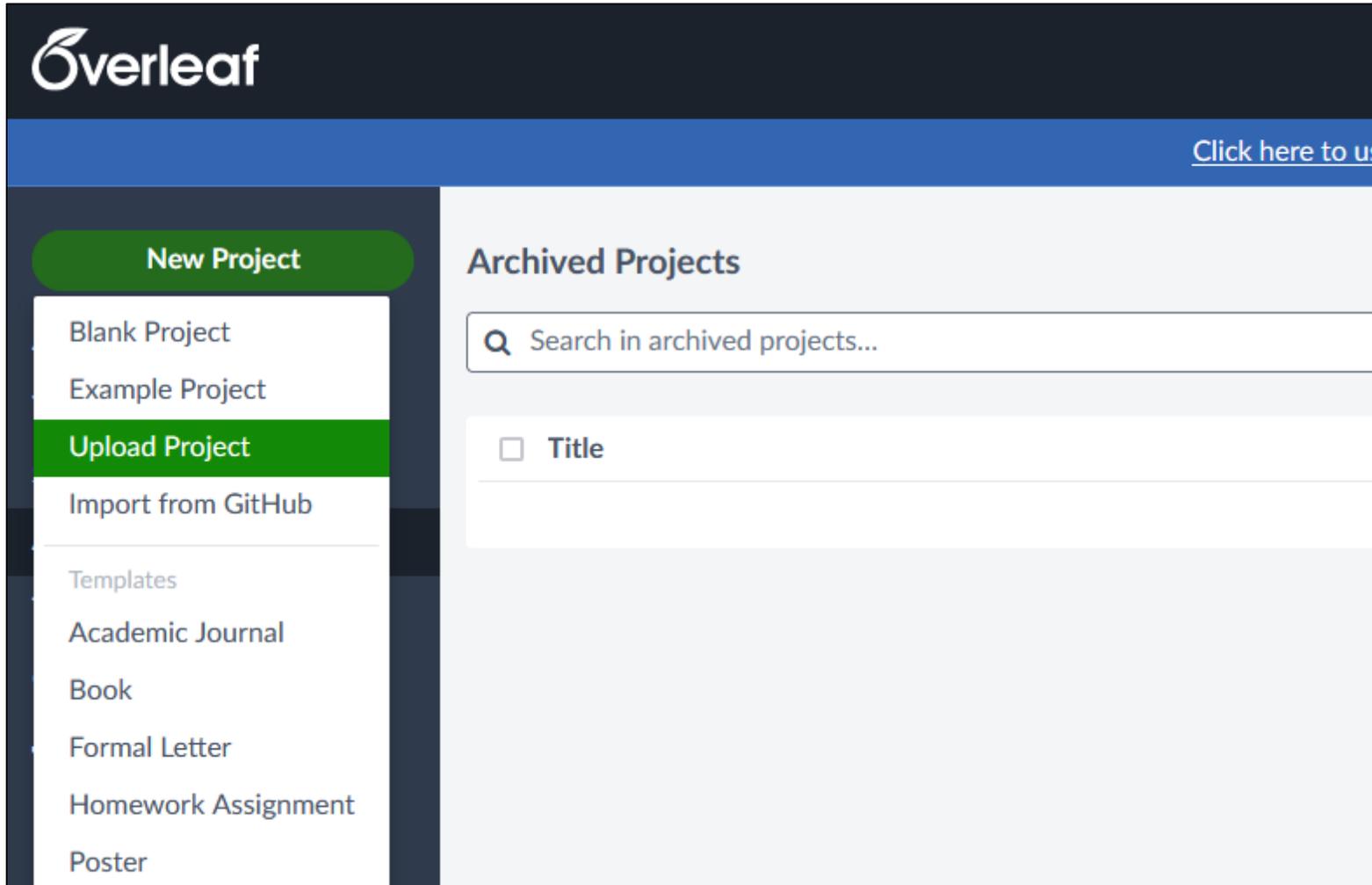
`\conclusion{Amoxicillin/potassium clavulanate results in significantly more cures and fewer failures than placebo}`.

```
\end{document}
```

Complete documentation and latest version:

<https://github.com/Christof93/SciKGT_EX>

Set up a LaTeX Project (using Overleaf)



1. Download Example.zip
<https://bit.ly/49RjRbO>
2. Open Overleaf
<https://www.overleaf.com/>
3. Upload Example.zip
New Project → Upload Project
4. Select or drag zip file

Remark:

We only use an abstract for the annotations as a simplified example. Annotations can be used **anywhere** (in one or more LaTeX files).

0. Add SciKGT_EX Files to the Project

The screenshot displays the Overleaf online LaTeX editor. The interface includes a top navigation bar with options like 'Menu', 'Upgrade', 'Review', 'Share', 'Submit', 'History', 'Layout', and 'Chat'. The main workspace is split into two panes:

- Left Pane (Code Editor):** Shows the LaTeX source code for 'example.tex'. The code includes:


```

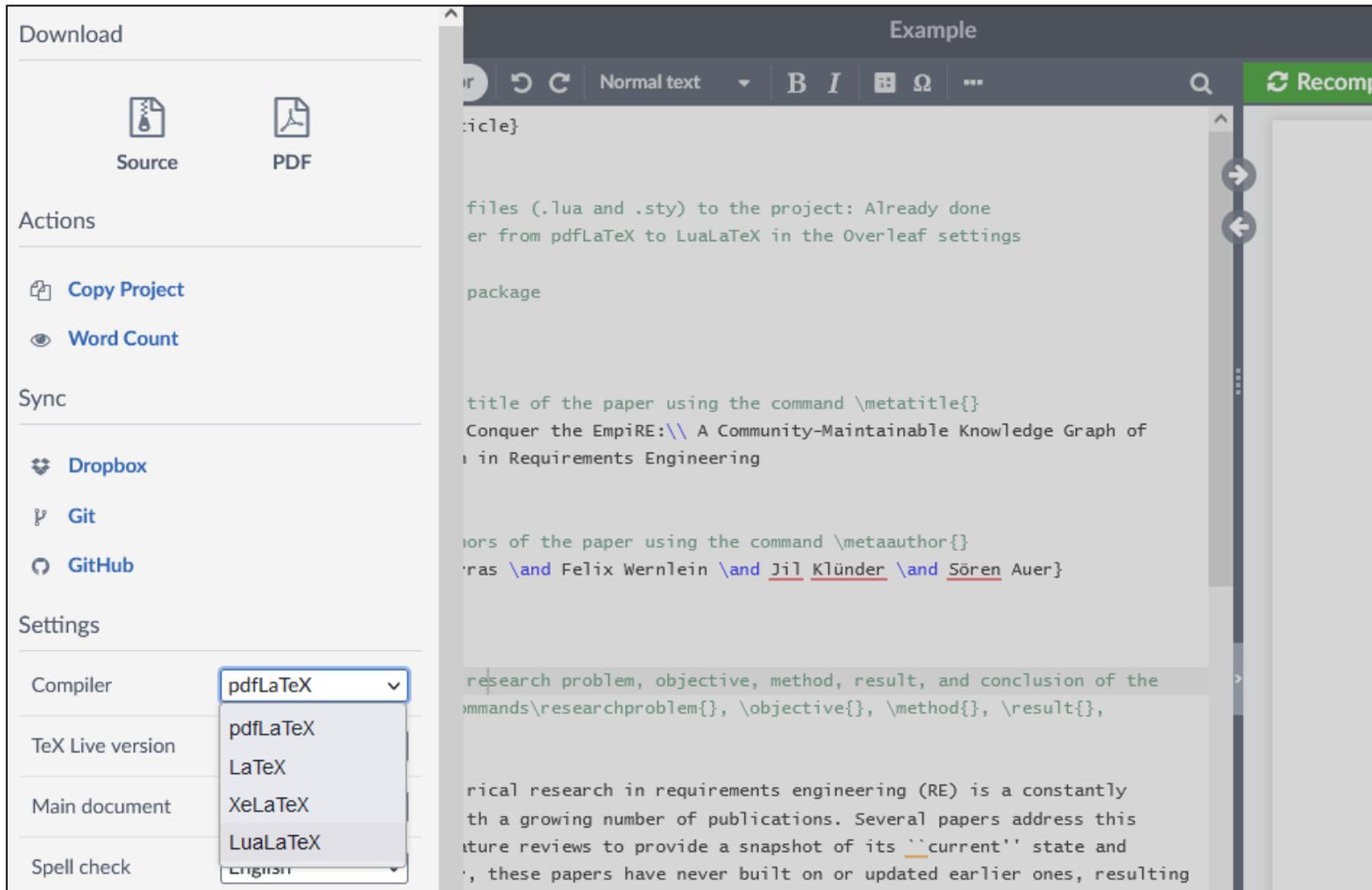
1 \documentclass{article}
2 \usepackage{ur1}
3
4 % 0. Add scikgtex files (.lua and .sty) to the project: Already done
5 % 1. Change compiler from pdfLaTeX to LuaLaTeX in the Overleaf settings
6
7 % 2. Add SciKGTEX package
8
9 \begin{document}
10
11 % 3. Annotate the title of the paper using the command \metatitle{}
12 \title{Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph
13 of Empirical Research in Requirements Engineering}
14 }
15 % 4. Annotate authors of the paper using the command \metaauthor{}
16 \author{Oliver Karras \and Felix Wernlein \and Jil Klünder \and Sören Auer}
17
18 \maketitle
19
20 % 5. Annotate the research problem, objective, method, result, and conclusion of
21 the paper using the commands \researchproblem{}, \objective{}, \method{},
22 \result{}, \conclusion{}
23
24 \begin{abstract}
25 [Background.] Empirical research in requirements engineering (RE) is a constantly
26 evolving topic, with a growing number of publications. Several papers address this
27 topic using literature reviews to provide a snapshot of its "current" state and
28 evolution. However, these papers have never built on or updated earlier ones,
29 resulting in overlap and redundancy. The underlying problem is the unavailability
30 of data from earlier works. Researchers need technical infrastructures to conduct
31 sustainable literature reviews. [Aims.] We examine the use of the Open Research
32 Knowledge Graph (ORKG) as such an infrastructure to build and publish an initial
33 Knowledge Graph of Empirical research in RE (KG-EmpIRE) whose data is openly
34 available. Our long-term goal is to continuously maintain KG-EmpIRE with the
35 research community to synthesize a comprehensive, up-to-date, and long-term
36 available overview of the state and evolution of empirical research in RE.
37 [Method.] We conduct a literature review using the ORKG to build and publish KG-
38 EmpIRE which we evaluate against competency questions derived from a published
39 vision of empirical research in software (requirements) engineering for 2020 --
40 2025. [Results.] From 570 papers of the IEEE International Requirements
41 Engineering Conference (2000 -- 2022), we extract and analyze data on the reported
42 empirical research and answer 16 out of 77 competency questions. These answers
43 show a positive development towards the vision, but also the need for future
44 sustainable literature reviews.
45 \end{abstract}
46 \end{document}

```
- Right Pane (Rendered PDF):** Shows the compiled document. The title is "Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering". The authors listed are Oliver Karras, Felix Wernlein, Jil Klünder, and Sören Auer. The date is April 18, 2024. The abstract text is visible below the title.

The following two files must be present:

- scikgtex.lua
- scikgtex.sty

1. Change Compiler from pdfLaTeX to LuaLaTeX



It is necessary to compile your LaTeX source with **LuaLaTeX** for the SciKGT_EX package to work.

2. Add the SciKGT_EX Package to the LaTeX File

```
1 \documentclass{article}
2 \usepackage{url}
3
4 % 0. Add scikgtex files (.lua and .sty) to the project: Already done
5 % 1. Change compiler from pdfLaTeX to LuaLaTeX in the Overleaf settings
6
7 % 2. Add SciKGTEX package
8
9
10 \begin{document}
11
```

Add the command to the preamble of the LaTeX file:

```
\usepackage{scikgtex}
```

2. Add the SciKGT_EX Package to the LaTeX File: Result

```
1 \documentclass{article}
2 \usepackage{url}
3
4 % 0. Add scikgtex files (.lua and .sty) to the project: Already done
5 % 1. Change compiler from pdfLaTeX to LuaLaTeX in the Overleaf settings
6
7 % 2. Add SciKGTEX package
8 \usepackage{scikgtex}
9
10 \begin{document}
11
```

Add the command to the preamble of the LaTeX file:

```
\usepackage{scikgtex}
```

3. & 4. Annotate Metadata of the Paper

```
10 ▾ \begin{document}
11
12 % 3. Annotate the title of the paper using the command \metatitle{}
13 ▾ \title{Divide and Conquer the EmpiRE:\\ A Community-Maintainable Knowledge Graph of
    Empirical Research in Requirements Engineering
14 }
15
16 % 4. Annotate authors of the paper using the command \metaauthor{}
17 \author{Oliver Karras \and Felix Wernlein \and Jil Klünder \and Sören Auer}
18
19 \maketitle
20
```

1. Annotate the title
`\metatitle{}`
2. Annotate the authors
`\metaauthor{}`

Remark:

The command `\metaauthor{}` must be used individually **for each author**. With four authors, you need the command four times.

3. & 4. Annotate Metadata of the Paper: Result

```
10 ▾ \begin{document}
11
12 % 3. Annotate the title of the paper using the command \metatitle{}
13 ▾ \title{\metatitle{Divide and Conquer the Empire:\\ A Community-Maintainable Knowledge
    Graph of Empirical Research in Requirements Engineering}
14 }
15
16 % 4. Annotate authors of the paper using the command \metaauthor{}
17 \author{\metaauthor{Oliver Karras} \and \metaauthor{Felix Wernlein} \and
    \metaauthor{Jil Klünder} \and \metaauthor{Sören Auer}}
18
19 \maketitle
20
```

1. Annotate the title
`\metatitle{}`
2. Annotate the authors
`\metaauthor{}`

Remark:

The command `\metaauthor{}` must be used individually **for each author**. With four authors, you need the command four times.

5. Annotate Content of the Paper

```
21 % 5. Annotate the research problem, objective, method, result, and conclusion of the
    paper using the commands\researchproblem{}, \objective{}, \method{}, \result{},
    \conclusion{}
22 \begin{abstract}
23 [Background.] Empirical research in requirements engineering (RE) is a constantly
    evolving topic, with a growing number of publications. Several papers address this
    topic using literature reviews to provide a snapshot of its ``current'' state and
    evolution. However, these papers have never built on or updated earlier ones, resulting
    in overlap and redundancy. The underlying problem is the unavailability of data from
    earlier works. Researchers need technical infrastructures to conduct sustainable
    literature reviews. [Aims.] We examine the use of the Open Research Knowledge Graph
    (ORKG) as such an infrastructure to build and publish an initial Knowledge Graph of
    Empirical research in RE (KG-EmpIRE) whose data is openly available. Our long-term goal
    is to continuously maintain KG-EmpIRE with the research community to synthesize a
    comprehensive, up-to-date, and long-term available overview of the state and evolution
    of empirical research in RE. [Method.] We conduct a literature review using the ORKG
    to build and publish KG-EmpIRE which we evaluate against competency questions derived from
    a published vision of empirical research in software (requirements) engineering for
    2020 -- 2025. [Results.] From 570 papers of the IEEE International Requirements
    Engineering Conference (2000 -- 2022), we extract and analyze data on the reported
    empirical research and answer 16 out of 77 competency questions. These answers show a
    positive development towards the vision, but also the need for future improvements.
    [Conclusions.] The ORKG is a ready-to-use and advanced infrastructure to organize data
    from literature reviews as knowledge graphs. The resulting knowledge graphs make the
    data openly available and maintainable by research communities, enabling sustainable
    literature reviews.
24 \end{abstract}
```

1. Annotate the research problem
`\researchproblem{}`
2. Annotate the objective
`\objective{}`
3. Annotate the method
`\method{}`
4. Annotate the result
`\result{}`
5. Annotate the conclusion
`\conclusion{}`

5. Annotate Content of the Paper

21 % 5. Annotate the research problem, objective, method, result, and conclusion of the

Remark:

- All commands can be used **multiple** times.
- The annotated text elements should be as **short** as **possible** and as **long** as **necessary**.
- Definitions:
 - **Research problem**: Issue or gap in existing knowledge addressed by the paper.
 - **Objective**: Goal that the paper aims to achieve.
 - **Method**: Systematic approach, technique, or action plan used in the paper to achieve a goal and result.
 - **Result**: Outcome from a systematic approach, technique, or action plan used in the paper.
 - **Conclusion**: Findings from the analysis of the research results in the paper.

24 \end{abstract}

1. Annotate the research problem
`\researchproblem{}`
2. Annotate the objective
`\objective{}`
3. Annotate the method
`\method{}`
4. Annotate the result
`\result{}`
5. Annotate the conclusion
`\conclusion{}`

5. Annotate Content of the Paper: Result

```
21 % 5. Annotate the research problem, objective, method, result, and conclusion of the
    paper using the commands\researchproblem{}, \objective{}, \method{}, \result{},
    \conclusion{}
22 \begin{abstract}
23 [Background.] Empirical research in requirements engineering (RE) is a constantly
    evolving topic, with a growing number of publications. Several papers address this
    topic using literature reviews to provide a snapshot of its ``current'' state and
    evolution. However, these papers have never built on or updated earlier ones, resulting
    in overlap and redundancy. The underlying problem is the
    \researchproblem{unavailability of data from earlier works}. Researchers need technical
    infrastructures to conduct sustainable literature reviews. [Aims.] We examine the
    \objective{use of the Open Research Knowledge Graph (ORKG) as such an infrastructure to
    build and publish an initial Knowledge Graph of Empirical research in RE (KG-EmpIRE)
    whose data is openly available}. Our long-term goal is to continuously maintain KG-
    EmpiRE with the research community to synthesize a comprehensive, up-to-date, and long-
    term available overview of the state and evolution of empirical research in RE.
    [Method.] We conduct a \method{literature review using the ORKG} to build and publish
    KG-EmpIRE which we \method{evaluate against competency questions} derived from a
    published vision of empirical research in software (requirements) engineering for 2020
    -- 2025. [Results.] \result{From 570 papers of the IEEE International Requirements
    Engineering Conference (2000 -- 2022), we extract and analyze data on the reported
    empirical research} and \result{answer 16 out of 77 competency questions}. These
    answers show a positive development towards the vision, but also the need for future
    improvements. [Conclusions.] \conclusion{The ORKG is a ready-to-use and advanced
    infrastructure to organize data from literature reviews as knowledge graphs}. The
    resulting knowledge graphs make the data openly available and maintainable by research
    communities, enabling sustainable literature reviews.
24 \end{abstract}
```

1. Annotate the research problem

`\researchproblem{}`

2. Annotate the objective

`\objective{}`

3. Annotate the method

`\method{}`

4. Annotate the result

`\result{}`

5. Annotate the conclusion

`\conclusion{}`

6. Annotate Content of the Paper with Invisible Markup

```
21 % 5. Annotate the research problem, objective, method, result, and conclusion of the
    paper using the commands\researchproblem{}, \objective{}, \method{}, \result{},
    \conclusion{}
22 \begin{abstract}
23 [Background.] Empirical research in requirements engineering (RE) is a constantly
    evolving topic, with a growing number of publications. Several papers address this
    topic using literature reviews to provide a snapshot of its ``current'' state and
    evolution. However, these papers have never built on or updated earlier ones, resulting
    in overlap and redundancy. The underlying problem is the
    \researchproblem{unavailability of data from earlier works}. Researchers need technical
    infrastructures to conduct sustainable literature reviews. [Aims.] We examine the
    \objective{use of the Open Research Knowledge Graph (ORKG) as such an infrastructure to
    build and publish an initial Knowledge Graph of Empirical research in RE (KG-EmpIRE)
    whose data is openly available}. Our long-term goal is to continuously maintain KG-
    EmpIRE with the research community to synthesize a comprehensive, up-to-date, and long-
    term available overview of the state and evolution of empirical research in RE.
    [Method.] We conduct a \method{literature review using the ORKG} to build and publish
    KG-EmpIRE which we \method{evaluate against competency questions} derived from a
    published vision of empirical research in software (requirements) engineering for 2020
    -- 2025. [Results.] \result{From 570 papers of the IEEE International Requirements
    Engineering Conference (2000 -- 2022), we extract and analyze data on the reported
    empirical research} and \result{answer 16 out of 77 competency questions}. These
    answers show a positive development towards the vision, but also the need for future
    improvements. [Conclusions.] \conclusion{The ORKG is a ready-to-use and advanced
    infrastructure to organize data from literature reviews as knowledge graphs}. The
    resulting knowledge graphs make the data openly available and maintainable by research
    communities, enabling sustainable literature reviews.
24 \end{abstract}
```

What can I do, if the text is not suitable for annotation?

Example:

“... \method{evaluate against competency questions} ...”

Solution:

\method*{evaluation against competency questions}

Remark:

This text is added to the PDF metadata, but **not rendered** in the text of the PDF.

6. Annotate Content of the Paper with Invisible Markup: Result

```

21 % 5. Annotate the research problem, objective, method, result, and conclusion of the
    paper using the commands\researchproblem{}, \objective{}, \method{}, \result{},
    \conclusion{}
22 \begin{abstract}
23 [Background.] Empirical research in requirements engineering (RE) is a constantly
    evolving topic, with a growing number of publications. Several papers address this
    topic using literature reviews to provide a snapshot of its ``current'' state and
    evolution. However, these papers have never built on or updated earlier ones, resulting
    in overlap and redundancy. The underlying problem is the unavailability of data from
    earlier works. Researchers need technical infrastructures to conduct sustainable
    literature reviews. [Aims.] We examine the \objective{use of the Open Research
    Knowledge Graph (ORKG) as such an infrastructure to build and publish an initial
    Knowledge Graph of Empirical research in RE (KG-EmpIRE) whose data is openly
    available}. Our long-term goal is to continuously maintain KG-EmpIRE with the research
    community to synthesize a comprehensive, up-to-date, and long-term available overview
    of the state and evolution of empirical research in RE. [Method.] We conduct a
    \method{literature review using the ORKG} to build and publish KG-EmpIRE which we
    evaluate against competency questions derived from a published vision of empirical
    research in software (requirements) engineering for 2020 -- 2025. [Results.]
    \result{From 570 papers of the IEEE International Requirements Engineering Conference
    (2000 -- 2022), we extract and analyze data on the reported empirical research} and
    \result{answer 16 out of 77 competency questions}. These answers show a positive
    development towards the vision, but also the need for future improvements.
    [Conclusions.] \conclusion{The ORKG is a ready-to-use and advanced infrastructure to
    organize data from literature reviews as knowledge graphs}. The resulting knowledge
    graphs make the data openly available and maintainable by research communities,
    enabling sustainable literature reviews.
24 \end{abstract}
25
26 % 6. If written text is not suitable for annotation, we can also annotate invisible
    text using the *-notation
27 \researchproblem*{unavailability of the extracted and analyzed data from literature
    reviews}
28 \method*{evaluation against competency questions}

```

What can I do, if the text is not suitable for annotation?

Example:

“... \method{evaluate against competency questions} ...”

Solution:

\method*{evaluation against competency questions}

Remark:

This text is added to the PDF metadata, but **not rendered** in the text of the PDF.

Comparison of Annotated Paper Versions

```

21 % 5. Annotate the research problem, objective, method, result, and conclusion of the
    paper using the commands\researchproblem{}, \objective{}, \method{}, \result{},
    \conclusion{}
22 \begin{abstract}
23 [Background.] Empirical research in requirements engineering (RE) is a constantly
    evolving topic, with a growing number of publications. Several papers address this
    topic using literature reviews to provide a snapshot of its ``current'' state and
    evolution. However, these papers have never built on or updated earlier ones, resulting
    in overlap and redundancy. The underlying problem is the
    \researchproblem{unavailability of data from earlier works}. Researchers need technical
    infrastructures to conduct sustainable literature reviews. [Aims.] We examine the
    \objective{use of the Open Research Knowledge Graph (ORKG) as such an infrastructure to
    build and publish an initial Knowledge Graph of Empirical research in RE (KG-EmpIRE)
    whose data is openly available}. Our long-term goal is to continuously maintain KG-
    EmpIRE with the research community to synthesize a comprehensive, up-to-date, and long-
    term available overview of the state and evolution of empirical research in RE.
    [Method.] We conduct a \method[literature review using the ORKG] to build and publish
    KG-EmpIRE which we \method{evaluate against competency questions} derived from a
    published vision of empirical research in software (requirements) engineering for 2020
    -- 2025. [Results.] \result{From 570 papers of the IEEE International Requirements
    Engineering Conference (2000 -- 2022), we extract and analyze data on the reported
    empirical research} and \result{answer 16 out of 77 competency questions}. These
    answers show a positive development towards the vision, but also the need for future
    improvements. [Conclusions.] \conclusion{The ORKG is a ready-to-use and advanced
    infrastructure to organize data from literature reviews as knowledge graphs}. The
    resulting knowledge graphs make the data openly available and maintainable by research
    communities, enabling sustainable literature reviews.
24 \end{abstract}

```

```

21 % 5. Annotate the research problem, objective, method, result, and conclusion of the
    paper using the commands\researchproblem{}, \objective{}, \method{}, \result{},
    \conclusion{}
22 \begin{abstract}
23 [Background.] Empirical research in requirements engineering (RE) is a constantly
    evolving topic, with a growing number of publications. Several papers address this
    topic using literature reviews to provide a snapshot of its ``current'' state and
    evolution. However, these papers have never built on or updated earlier ones, resulting
    in overlap and redundancy. The underlying problem is the unavailability of data from
    earlier works. Researchers need technical infrastructures to conduct sustainable
    literature reviews. [Aims.] We examine the \objective{use of the Open Research
    Knowledge Graph (ORKG) as such an infrastructure to build and publish an initial
    Knowledge Graph of Empirical research in RE (KG-EmpIRE) whose data is openly
    available}. Our long-term goal is to continuously maintain KG-EmpIRE with the research
    community to synthesize a comprehensive, up-to-date, and long-term available overview
    of the state and evolution of empirical research in RE. [Method.] We conduct a
    \method[literature review using the ORKG] to build and publish KG-EmpIRE which we
    evaluate against competency questions derived from a published vision of empirical
    research in software (requirements) engineering for 2020 -- 2025. [Results.]
    \result{From 570 papers of the IEEE International Requirements Engineering Conference
    (2000 -- 2022), we extract and analyze data on the reported empirical research} and
    \result{answer 16 out of 77 competency questions}. These answers show a positive
    development towards the vision, but also the need for future improvements.
    [Conclusions.] \conclusion{The ORKG is a ready-to-use and advanced infrastructure to
    organize data from literature reviews as knowledge graphs}. The resulting knowledge
    graphs make the data openly available and maintainable by research communities,
    enabling sustainable literature reviews.
24 \end{abstract}
25
26 % 6. If written text is not suitable for annotation, we can also annotate invisible
    text using the *-notation
27 \researchproblem*{unavailability of the extracted and analyzed data from literature
    reviews}
28 \method*{evaluation against competency questions}

```

7. (Optional) Annotate Research Field

Remark:

1. A paper in the ORKG is assigned to a research field based on the DFG classification. All research fields: <https://orkg.org/fields>.
2. URI of the ORKG semantic web resource for Software Engineering: <https://orkg.org/resource/R140>.
3. Use the command `\uri{"URI"}{"Label"}` inside an annotation to refer to resources in the semantic web. The first argument is the **URI** to the semantic resource and the second is an optional **Label**.
4. The term "Software Engineering" does not appear in the abstract, so we need an **invisible** annotation and we also use the `\uri{}` command to create a reference to the semantic resource.

1. Annotate the research field

`\researchfield{}`

2. Refer to the ORKG semantic web resource [Software Engineering](https://orkg.org/resource/R140)

`\uri{"URI"}{"Label"}`

7. (Optional) Annotate Research Field: Result

```
29 % 7. Optional: Annotate research field of the paper
30 \researchfield*{\uri{https://orkg.org/resource/R140}{Software Engineering}}
```

Remark:

1. A paper in the ORKG is assigned to a research field based on the DFG classification. All research fields: <https://orkg.org/fields>.
2. URI of the ORKG semantic web resource for Software Engineering: <https://orkg.org/resource/R140>.
3. Use the command `\uri{“URI”}{“Label”}` inside an annotation to refer to resources in the semantic web. The first argument is the **URI** to the semantic resource and the second is an optional **Label**.
4. The term “Software Engineering” does not appear in the abstract, so we need an **invisible** annotation and we also use the `\uri{}` command to create a reference to the semantic resource.

1. Annotate the research field

`\researchfield{}`

2. Refer to the ORKG semantic web resource [Software Engineering](https://orkg.org/resource/R140)

`\uri{“URI”}{“Label”}`

8. Use Custom Annotations of REFSQ'24 and REFSQ'25

```
33 % 8. Optional: Using REFSQ'24 and REFSQ'25 annotations
34 \contribution*{code repository}{\url{https://github.com/okarras/EmpIRE-Analysis}}
35 \contribution*{dataset}{\url{https://orkg.org/api/rdf/dump}}
36
37 \end{document}
```

Remark:

1. Use the command `\contribution{"Property name"}{"Label"}` to add a custom annotation for your domain. The first argument is the **Property name** of the property from the ORKG you want to use, and the second is an optional **Label**. SciKGT_{EX} checks if a property with the provided exact **Property name** exists and replaces it with the internal property ID in the ORKG namespace.
2. All ORKG properties can be found here: <https://orkg.org/properties>.

1. Annotate the code repository

`\contribution{code repository}{}`

2. Annotate the dataset

`\contribution{dataset}{}`

9. Generate FAIR-Annotated PDF of the Paper

The screenshot shows a LaTeX editor interface with a code editor on the left and a preview window on the right. The code editor contains the following LaTeX source code:

```

1 \documentclass{article}
2 \usepackage{ur}
3
4 % 0. Add scikgtex files (.lua and .sty) to the project: Already done
5 % 1. Change compiler from pdfLaTeX to LuaLaTeX in the Overleaf settings
6
7 % 2. Add ScikGTEx package
8 \usepackage{scikgtex}
9
10 \begin{document}
11
12 % 3. Annotate the title of the paper using the command \metatitle{}
13 \title{\metatitle{Divide and Conquer the EmpiRE: A Community-Maintainable
14 Knowledge Graph of Empirical Research in Requirements Engineering}
15 }
16
17 % 4. Annotate authors of the paper using the command \metaauthor{}
18 \author{\metaauthor{Oliver Karras} \and \metaauthor{Felix Wernlein} \and
19 \metaauthor{Jil Klünder} \and \metaauthor{Sören Auer}}
20
21 \maketitle
22
23 % 5. Annotate the research problem, objective, method, result, and conclusion of
24 the paper using the commands \researchproblem{}, \objective{}, \method{},
25 \result{}, \conclusion{}
26 \begin{abstract}
27 [Background.] Empirical research in requirements engineering (RE) is a constantly
28 evolving topic, with a growing number of publications. Several papers address this
29 topic using literature reviews to provide a snapshot of its "current" state and
30 evolution. However, these papers have never built on or updated earlier ones,
31 resulting in overlap and redundancy. The underlying problem is the unavailability
32 of data from earlier works. Researchers need technical infrastructures to conduct
33 sustainable literature reviews. [Aims.] We examine the objective{use of the Open
34 Research Knowledge Graph (ORKG) as such an infrastructure to build and publish an
35 initial Knowledge Graph of Empirical research in RE (KG-EmpIRE) whose data is
36 openly available}. Our long-term goal is to continuously maintain KG-EmpIRE with
37 the research community to synthesize a comprehensive, up-to-date, and long-term
38 available overview of the state and evolution of empirical research in RE.
39 [Method.] We conduct a \method{literature review using the ORKG} to build and
40 publish KG-EmpIRE which we evaluate against competency questions derived from a
41 published vision of empirical research in software (requirements) engineering for
42 2020 -- 2025. [Results.] \result{From 570 papers of the IEEE International
43 Requirements Engineering Conference (2000 -- 2022), we extract and analyze data on

```

The preview window shows the rendered PDF with the following content:

Divide and Conquer the EmpiRE:
A Community-Maintainable Knowledge Graph of
Empirical Research in Requirements Engineering

Oliver Karras Felix Wernlein Jil Klünder Sören Auer

April 19, 2024

Abstract

[Background.] Empirical research in requirements engineering (RE) is a constantly evolving topic, with a growing number of publications. Several papers address this topic using literature reviews to provide a snapshot of its "current" state and evolution. However, these papers have never built on or updated earlier ones, resulting in overlap and redundancy. The underlying problem is the unavailability of data from earlier works. Researchers need technical infrastructures to conduct sustainable literature reviews. [Aims.] We examine the use of the Open Research Knowledge Graph (ORKG) as such an infrastructure to build and publish an initial Knowledge Graph of Empirical research in RE (KG-EmpIRE) whose data is openly available. Our long-term goal is to continuously maintain KG-EmpIRE with the research community to synthesize a comprehensive, up-to-date, and long-term available overview of the state and evolution of empirical research in RE. [Method.] We conduct a literature review using the ORKG to build and publish KG-EmpIRE which we evaluate against competency questions derived from a published vision of empirical research in software (requirements) engineering for 2020 -- 2025. [Results.] From 570 papers of the IEEE International Requirements Engineering Conference (2000 -- 2022), we extract and analyze data on the reported empirical research and answer 16 out of 77 competency questions. These answers show a positive development towards the vision, but also the need for future improvements. [Conclusions.] The ORKG is a ready-to-use and advanced infrastructure to organize data from literature reviews as knowledge graphs. The resulting knowledge graphs make the data openly available and maintainable by research communities, enabling sustainable literature reviews.

Simply...
recompile

Remark:
Each recompile
adds the
annotations to the
metadata of the
generated PDF.

9. Generate FAIR-Annotated PDF of the Paper

The screenshot shows a LaTeX editor interface with a document being compiled. The document content includes:

```

1 \documentclass{article}
2 \usepackage{ur}
3
4 % 0. Add scikgtex files (.lua and .sty) to the project: Already done
5 % 1. Change compiler from pdfLaTeX to LuaLaTeX in the Overleaf settings
6
7 % 2. Add SciKGTex package
8 \usepackage{scikgtex}
9
10 \begin{document}
11
12 % 3. Annotate the title of the paper using the command \metatitle{}
13 \title{\metatitle{Divide and Conquer the EMPIRE: A Community-Maintainable
14 Knowledge Graph of Empirical Research in Requirements Engineering}}
15
16 % 4. Annotate authors of the paper using the command \metaauthor{}
17 \author{\metaauthor{Oliver Karras} \and \metaauthor{Felix Wernlein} \and
18 \metaauthor{Jiri Klunder} \and \metaauthor{Soren Auer}}
19
20 \maketitle
21
22 % 5. Annotate the research problem, objective, method, result, and conclusion of
23 the paper using the commands \researchproblem{}, \objective{}, \method{},
24 \result{}, \conclusion{}
25 \begin{abstract}
26 [Background.] Empirical research in requirements engineering (RE) is a constantly
27 evolving topic, with a growing number of publications. Several papers address this
28 topic using literature reviews to provide a snapshot of its ``current'' state and
29 evolution. However, these papers have never built on or updated earlier ones,
30 resulting in overlap and redundancy. The underlying problem is the unavailability
31 of data from earlier works. Researchers need technical infrastructures to conduct
32 sustainable literature reviews. [Aims.] We examine the use of the Open Research
33 Knowledge Graph (ORKG) as such an infrastructure to build and publish an initial
34 Knowledge Graph of Empirical research in RE (KG-EMPIRE) whose data is openly
35 available. Our long-term goal is to continuously maintain KG-EMPIRE with the
36 research community to synthesize a comprehensive, up-to-date, and long-term
37 available overview of the state and evolution of empirical research in RE.
38 [Method.] We conduct a literature review using the ORKG to build and publish KG-
39 EMPIRE which we evaluate against competency questions derived from a published
40 vision of empirical research in software (requirements) engineering for 2020 --
41 2025. [Results.] From 570 papers of the IEEE International Requirements
42 Engineering Conference (2000 -- 2022), we extract and analyze data on the reported

```

On the right side of the editor, there are five orange warning messages:

- Package SciKGTex Warning: No researchproblem annotation found! Are you sure you don't want to mark an entity with researchproblem?
- Package SciKGTex Warning: No objective annotation found! Are you sure you don't want to mark an entity with objective?
- Package SciKGTex Warning: No method annotation found! Are you sure you don't want to mark an entity with method?
- Package SciKGTex Warning: No result annotation found! Are you sure you don't want to mark an entity with result?
- Package SciKGTex Warning: No conclusion annotation found! Are you sure you don't want to mark an entity with conclusion?

Below the warnings is a "Raw logs" section showing the compilation output:

```

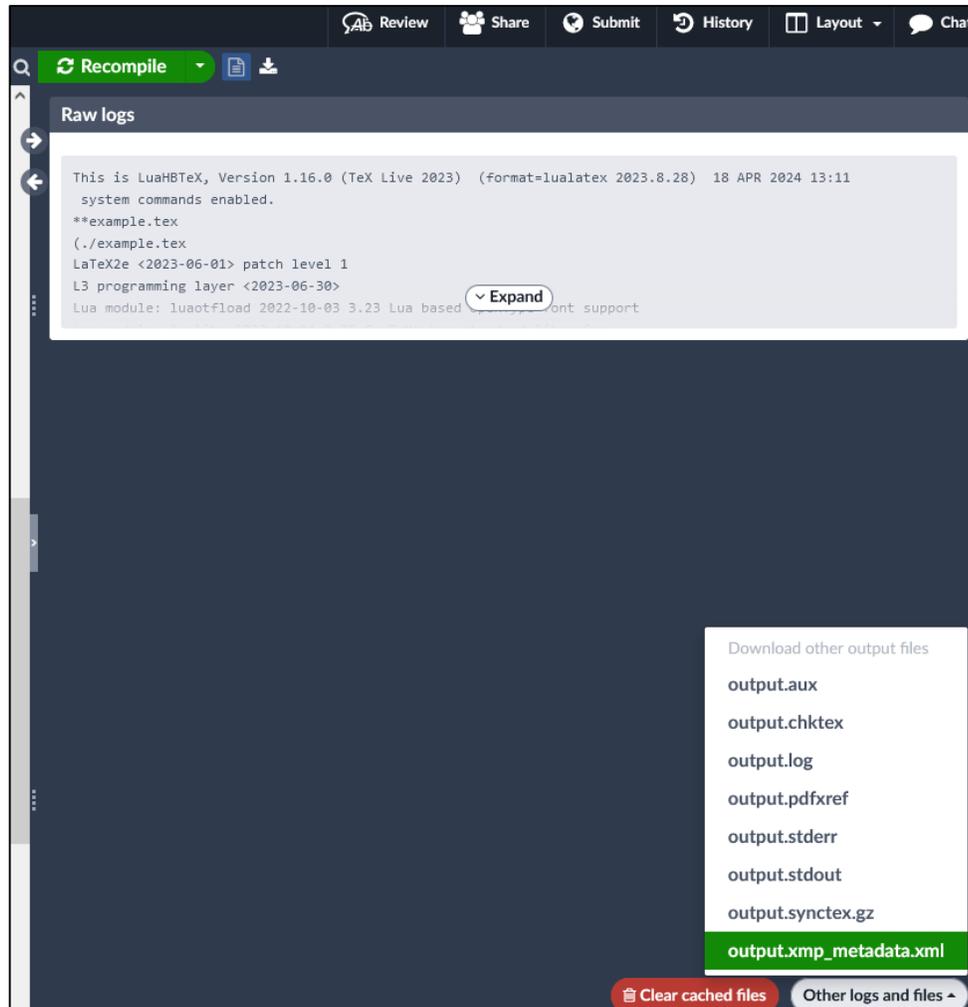
This is LuaHBTeX, Version 1.16.0 (TeX Live 2023) (format=lualatex 2023.8.28) 19 APR 2024 06:36
system commands enabled.
**example.tex
(./example.tex
LaTeX2e <2023-06-01> patch level 1
L3 programming layer <2023-06-30>
Lua module: luaotfload 2022-10-03 3.23 Lua base font support

```

At the bottom right, there are buttons for "Clear cached files" and "Other logs and files".

For the 5 predefined commands for content, SciKGTex also provides **warnings** if an annotation is missing.

10. Check the FAIR information embedded in the PDF



1. Open “Logs and outputs files”
2. Select “Other logs and files”
3. Download “output.xmp_metadata.xml”
4. Open “output.xmp_metadata.xml”

Remark:

We only use an abstract for the annotations as a simplified example. Annotations can be used **anywhere** (in one or more LaTeX files).

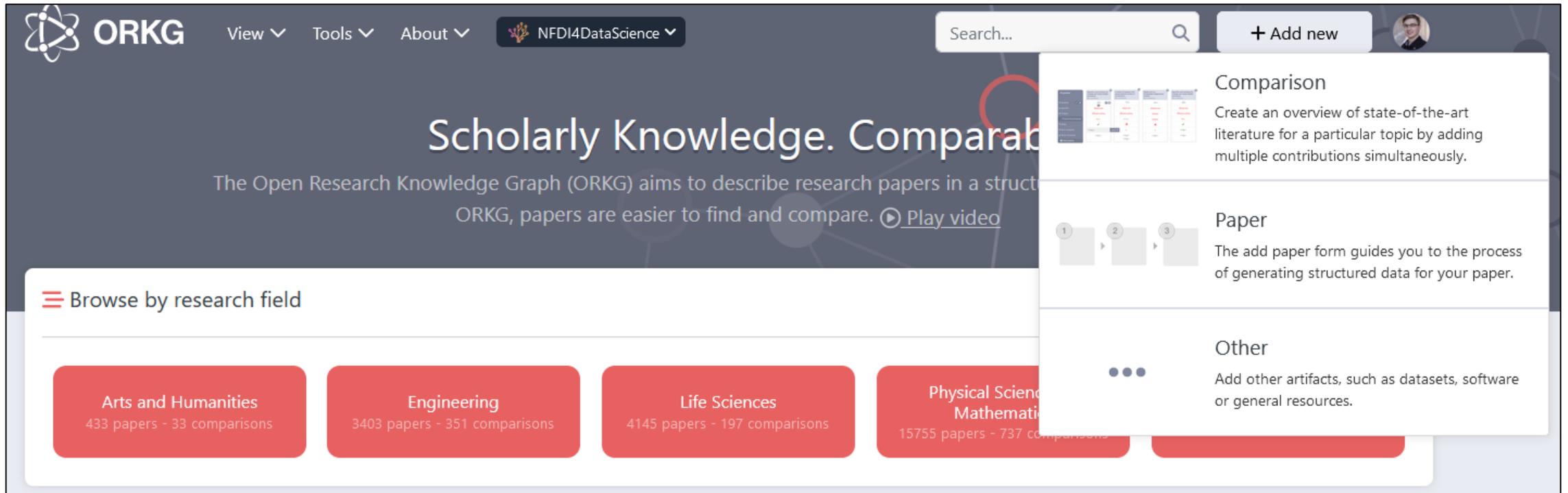
10. Check the FAIR information embedded in the PDF

```

<?xpacket begin="?" id="a7c48312-233b-400b-c0f3-0c296941c6"?>
▼<x:xmpmeta xmlns:x="adobe:ns:meta/">
  ▼<rdf:RDF xmlns:orkg="http://orkg.org/core#" xmlns:orkg_property="http://orkg.org/property/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    ▼<rdf:Description rdf:about="https://www.orkg.org/orkg/paper/a7c48312-233b-400b-c0f3-0c296941c6">
      <rdf:type rdf:resource="http://orkg.org/core#Paper"/>
      <orkg:hasTitle>Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements
      Engineering</orkg:hasTitle>
      <orkg:hasAuthor>Oliver Karras</orkg:hasAuthor>
      <orkg:hasAuthor>Felix Wernlein</orkg:hasAuthor>
      <orkg:hasAuthor>Jil Klünder</orkg:hasAuthor>
      <orkg:hasAuthor>Sören Auer</orkg:hasAuthor>
    ▼<orkg_property:P30>
      ▼<rdf:Description rdf:about="https://orkg.org/resource/R140">
        <rdfs:label>Software Engineering</rdfs:label>
        </rdf:Description>
      </orkg_property:P30>
    ▼<orkg:hasResearchContribution>
      ▼<orkg:ResearchContribution rdf:about="https://www.orkg.org/orkg/paper/a7c48312-233b-400b-c0f3-0c296941c6/contribution_ORKG_default">
        <orkg_property:P15051>use of the Open Research Knowledge Graph (ORKG) as such an infrastructure to build and publish an initial Knowledge
        Graph of Empirical research in RE (KG-EmpIRE) whose data is openly available</orkg_property:P15051>
        <orkg_property:P1005>literature review using the ORKG</orkg_property:P1005>
        <orkg_property:P1006>From 570 papers of the IEEE International Requirements Engineering Conference (2000 -- 2022), we extract and analyze
        data on the reported empirical research</orkg_property:P1006>
        <orkg_property:P1006>answer 16 out of 77 competency questions</orkg_property:P1006>
        <orkg_property:P15419>The ORKG is a ready-to-use and advanced infrastructure to organize data from literature reviews as knowledge
        graphs</orkg_property:P15419>
        <orkg_property:P32>unavailability of the extracted and analyzed data from literature reviews</orkg_property:P32>
        <orkg_property:P1005>evaluation against competency questions</orkg_property:P1005>
        <orkg_property:P49000>https://github.com/okarras/EmpIRE-Analysis</orkg_property:P49000>
        <orkg_property:P2005>https://orkg.org/api/rdf/dump</orkg_property:P2005>
      </orkg:ResearchContribution>
    </orkg:hasResearchContribution>
  </rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end="r"?>

```

11. Import the FAIR information into the ORKG



The screenshot shows the ORKG website interface. At the top left is the ORKG logo and navigation links: View, Tools, About, and a user profile dropdown for 'NFDI4DataScience'. A search bar and a '+ Add new' button are on the top right. The main heading is 'Scholarly Knowledge. Comparison'. Below it, a sub-heading reads 'The Open Research Knowledge Graph (ORKG) aims to describe research papers in a structured way. In ORKG, papers are easier to find and compare.' A 'Play video' button is also present. A 'Browse by research field' section contains several red buttons with the following data:

Research Field	Papers	Comparisons
Arts and Humanities	433	33
Engineering	3403	351
Life Sciences	4145	197
Physical Sciences and Mathematics	15755	737

The '+ Add new' dropdown menu is open, showing three options: 'Comparison' (with a thumbnail of a comparison interface), 'Paper' (with a thumbnail of a paper form), and 'Other' (with a thumbnail of three dots). The 'Paper' option is highlighted with a red circle in the original image.

1. Log in with your credentials
2. Click “+ Add new”
3. Select “Paper”

Remark:

ORKG does **not save** the uploaded PDF. It **only extracts** the FAIR annotations embedded in the PDF metadata.

11. Import the FAIR information into the ORKG

ORKG View Tools About NFDI4DataScience Search... + Add new

Add paper Upload PDF Enter BibTeX

DOI ?

Lookup

When a DOI is entered, some metadata is automatically filled

Show metadata fields [Click here if you don't have a DOI](#)

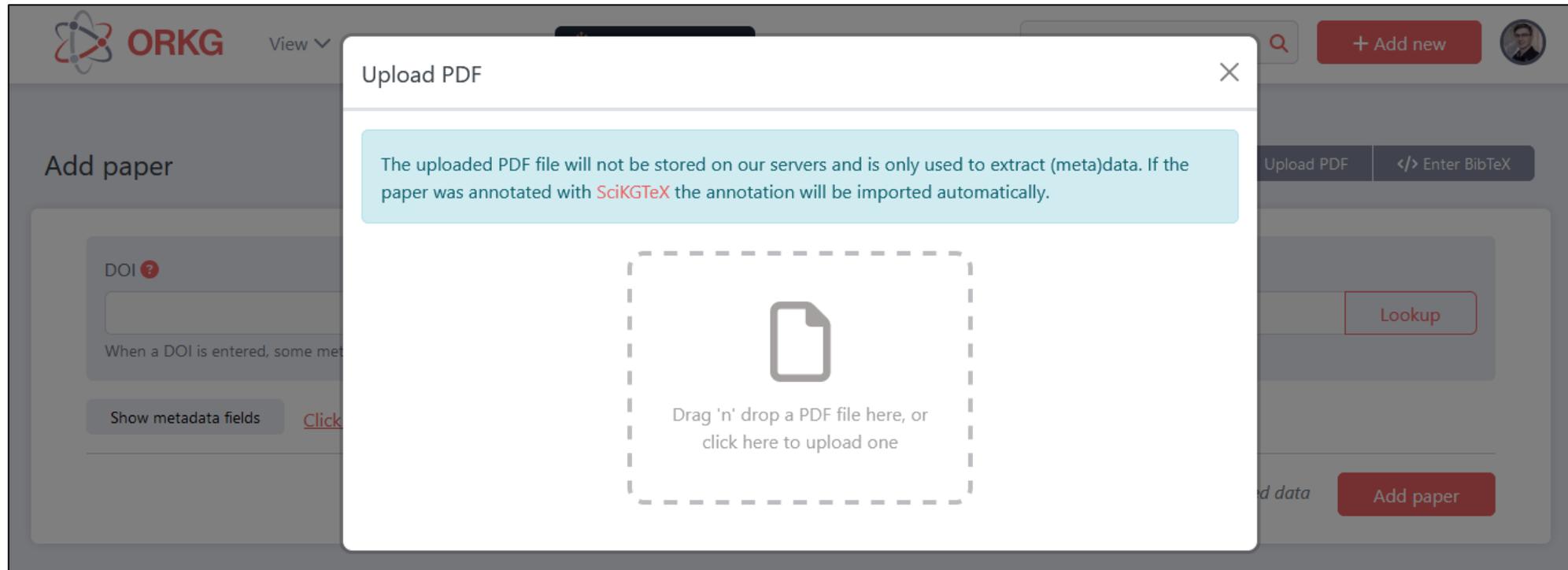
After adding the paper, you will be able to add structured data Add paper

1. Select “Upload PDF”

Remark:

ORKG does **not save** the uploaded PDF. It **only extracts** the FAIR annotations embedded in the PDF metadata.

11. Import the FAIR information into the ORKG



1. Select or drag your PDF file

Remark:

ORKG does **not save** the uploaded PDF. It **only extracts** the FAIR annotations embedded in the PDF metadata.

11. Import the FAIR information into the ORKG

ORKG View Tools About NFDI4DataScience Search... + Add new

Add paper Upload PDF Enter BibTeX

DOI Lookup
When a DOI is entered, some metadata is automatically filled

Hide metadata fields

Paper title (required)

Research field (required) Choose

Paper authors

- Oliver Karras
- Felix Wernlein
- Jil Klünder
- Sören Auer

+ Add author

Publication month Publication year

Published in

Paper URL

After adding the paper, you will be able to add structured data

1. ORKG shows imported metadata
2. Optional: Edit the metadata
3. Select “Add paper” at the bottom

Remark:

ORKG does **not save** the uploaded PDF. It **only extracts** the FAIR annotations embedded in the PDF metadata.

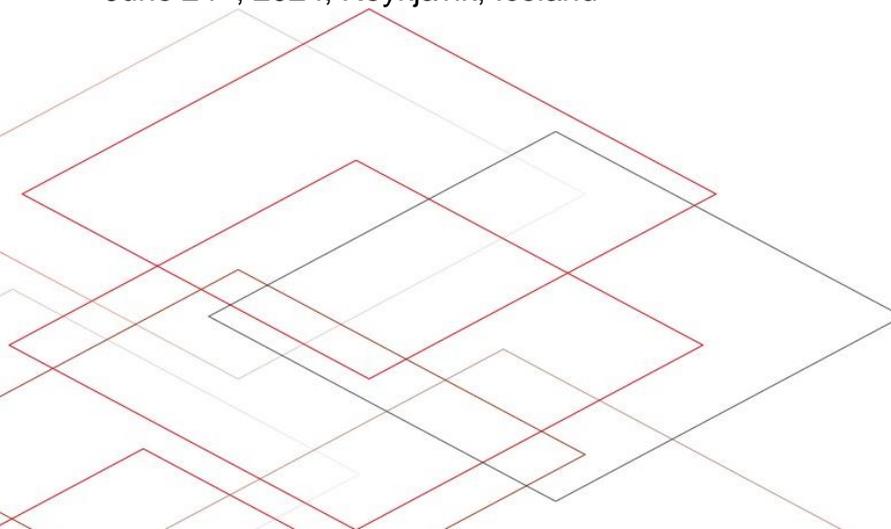
ORKG: Describing Papers Manually and Creating an ORKG Comparison

Oliver Karras, Alessio Ferrari, Davide Fucci, and Davide Dell'Anna

oliver.karras@tib.eu, alessio.ferrari@isti.cnr.it, davide.fucci@bth.se, d.dellanna@uu.nl

32nd IEEE International Requirements Engineering 2024 Conference – Exploring New Horizons: Expanding the Frontiers of Requirements Engineering

June 24th, 2024, Reykjavik, Iceland



ORKG Comparisons

Acknowledgement
of creators

DOI

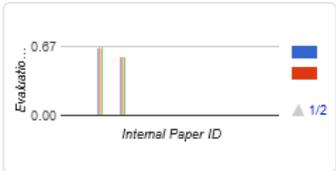
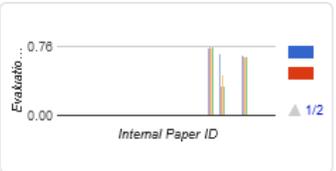
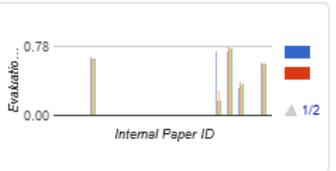
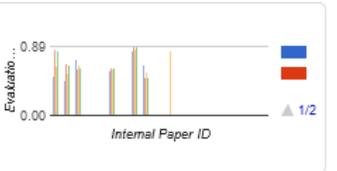
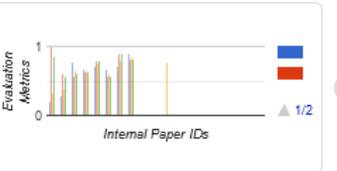
Visualizations

Interactive filtering

Overview of Approaches that Classify User Feedback as Feature Request ☆ 👁

📅 June 2021 👤 Oliver Karras 👤 Eduard C. Groen

This overview shows the classification results of approaches that use the machine learning algorithms Naïve Bayes, Support Vector Machines, and Decision Trees C4.5 in combination with the machine learning features Bag of Words or Term Frequency - Inverse Document Frequency to classify user feedback as feature request.

Properties	Software Feature Request Detection in Issue Tracking Systems <i>User Feedback Classification - 2016</i>	Mining User Requirements from Application Store Reviews Using Frame Semantics <i>User Feedback Classification - 2017</i>	Mining Twitter Feeds for Software User Requirements <i>User Feedback Classification - 2017</i>	Automatic Classification of Non-Functional Requirements from Augmented App User Reviews <i>User Feedback Classification - 2017</i>	Bug reports simply classified
has dataset	https://zenodo.org/record/56907#.YKT_NudCRPY	https://mast.informatik.uni-hamburg.de/wp-content/uploads/2014/03/REJ_data.zip https://sites.google.com/site/appsuserreviews/ seel.cse.lsu.edu/data/refsq17.zip	seel.cse.lsu.edu/data/re17.zip	Not available	https://seel.cse.lsu.edu/data/re17.zip

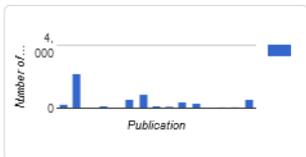
Example: ORKG Comparison of Related Work – State of the Art^[5]

A Comparison of Scientific Publications on the State of Empirical Research in Requirements Engineering and Software Engineering

November 2023 | Oliver Karras | Felix Wernlein | Jil Ann-Christin Klünder | Sören Auer

This comparison provides an overview of scientific publications that have investigated primary studies in requirements engineering and software engineering to give a snapshot of the "current" state of empirical research in requirements engineering and software engineering. In particular, the comparison shows for each publication (1) which research fields and topics were investigated, (2) whether and where the extracted and analyzed data is available, and (3) which method was used to determine the state, including further details about the respective method.

DOI: <https://doi.org/10.48366/R650023>



Properties	Status of Empirical Research in Software Engineering Empirical research - 2007	The type of evidence produced by empirical software engineers Empirical research - 2005	Research in software engineering: an analysis of the literature Empirical research - 2002
method	Literature review	Literature review	Literature review
data availability	×	×	×
time interval/era covered	1996-01-01	1997-01-01	1999-01-01
"last harvested"	2006-09-30	2003-12-31	1999-12-31
number of papers	113	119	360

Figure 1: Comparison of related publications on the "current" state and evolution of empirical research in RE and SE [16].

Figure 1 shows an excerpt from a comparison that we created to get an overview of related publications on the "current" state and evolution of empirical research in RE and SE [16]. For three publications, the excerpt shows the method used, the data availability, as well as the period and the number of papers examined. We use the ORKG due to its cross-domain and cross-topic characteristics, as well as its successful application for CrowdRE by Karras et al. [42].

III. RELATED WORK

Below, we review 14 publications that provide snapshots of the "current" state and evolution of empirical research in RE and SE (see Table I) [16]. We only consider publications that address the topic in general and are not limited to specific aspects, such as a method [67], [68] or a context [69], [70].

We found five publications on empirical research in RE published between 2005 and 2016 and nine on empirical research in SE published between 2002 and 2021. While one publication [1] examined empirical research in RE using a survey with 42 respondents, the other 13 publications [2], [3], [5]–[15] used (systematic) literature reviews or systematic mapping studies to analyze on average 402.9 papers (minimum: 20, median: 154, and maximum: 2237 papers) published between 1977 and 2019 with overlapping periods. In total, these 13 publications examined papers from a total of 60 different venues on 18 different themes. Nine of the 60 venues and ten of the 18 themes were examined by more than two publications. These facts show that there is considerable overlap and redundancy between these publications in terms of

goals, methods used, periods, venues, and themes examined. This overlap and redundancy could have been avoided if researchers had collaborated to build on and update earlier works. However, only four out of 14 publications offer their data at all, with only one publication [15] using a public data repository [71], [72]. The other three publications only offer links that no longer work [5], [12], [13].

In terms of key findings, the 14 publications show consistent results, although not all 18 themes were examined in all publications. For example, eleven of the 14 publications reported on the most commonly used research methods. Until 2000, the most common research methods were conceptual analysis and concept implementation [7]. Between 2000 and 2015, the most commonly used research methods changed to case studies and experiments [3], [5], [8]–[11], which were expanded after 2015 to also include surveys and systematic literature reviews [12]–[15]. While this change shows an evolution of research methods used, we also note that experiments and case studies have been the two main research methods for empirical research in RE and SE for more than 20 years. Although these two research methods have been used for a long time, seven publications concluded that there is a need to develop, expand, and use standardized terminology and theories (from other disciplines) to more consistently represent the empirical research conducted and better explain the results found [1], [2], [6]–[8], [12], [13]. In this regard, seven publications also analyzed the information reported for a comprehensive description of a research design. This information includes details about the research question(s) [1], contextual factors [6],

Table I: Details of related publications on the "current" state and evolution of empirical research in RE and SE [16]. Legend: Literature Review (LR), Systematic Literature Review (SLR), and Systematic Mapping Study (SMS)

Paper	Year	Field	Method	Period	Data basis	Dataset	Venues (Frequency > 2)	Themes (Frequency > 2)
[6]	2005	RE	LR	1968 – 2002	35 papers	Unavailable		
[5]	2010	RE	SLR	1977 – 2009	154 papers	Unavailable		
[11]	2012	RE	Survey	1970.03.2012 – 2011	42 respondents	Unavailable	1) Empirical Software Engineering Journal (8)	1) Data collection (12)
[2]	2014	RE	LR	2005 – 2011	2237 papers	Unavailable	2) IEEE Software (4)	2) Research method (11)
[3]	2016	SE	SMS	Open – 2012	290 papers	Unavailable	3) Requirements Engineering Journal (4)	3) Bibliographic metadata (10)
[8]	2002	SE	LR	1997 – 1998	707 papers	Unavailable	4) ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (4)	4) Data analysis (8)
[12]	2002	SE	LR	1996 – 2002	68 papers	Unavailable	5) IEEE Transactions on Software Engineering (3)	5) Research paradigms (7)
[1]	2005	RE	LR	1997 – 2005	119 papers	Unavailable	6) International and Software Technology Journal (3)	6) Research design (7)
[10]	2006	SE	LR	1997 – 2005	65 papers	Unavailable	7) IEEE International Requirements Engineering Conference (3)	7) Research topic (5)
[11]	2007	SE	LR	1996 – 2003	133 papers	Unavailable	8) Journal of Systems and Software (3)	8) Research context (4)
[14]	2015	SE	SMS	1996 – 2011	392 papers	Broken link	9) International Conference on Software Engineering (3)	9) Sample of population (4)
[13]	2018	SE	SMS	2017 – 2017	336 papers	Broken link		10) Theory (5)
[14]	2019	SE	SMS	2001 – 2014	343 papers	Unavailable		
[15]	2021	SE	SMS	Open – 2019	20 papers	Available		

Properties	Empirical research in requirements engineering: trends and opportunities <i>Empirical research - 2016</i>	Empirical research methodologies and studies in Requirements Engineering: How far did we come? <i>Empirical research - 2014</i>	A Survey on Empirical Requirements Engineering Research Practices <i>Empirical research - 2012</i>	Evidence-Based Structuring and Evaluation of Empirical Research in Requirements Engineering: Fundamentals, Framework, Research Map <i>Empirical research - 2010</i>	An Analy... <i>Empirical...</i>
research problem	empirical research in requirements engineering	empirical research in requirements engineering	empirical research in requirements engineering	empirical research in requirements engineering	empirical...
research field investigated	Requirements Engineering	Requirements Engineering	Requirements Engineering	Requirements Engineering	Requ...
topic investigated	bibliographic metadata context data collection	bibliographic metadata research topic theory	context	context	

<https://doi.org/10.48366/R650023>

[5] Karras et al.: *Divide and Conquer the EmpiRE: A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering*. 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), DOI: [10.1109/ESEM56168.2023.10304795](https://doi.org/10.1109/ESEM56168.2023.10304795), 2023.

Example: ORKG Comparison of Literature – ML Approaches^[9]

Overview of Approaches that Classify User Feedback as Feature Request ☆🔍

June 2021 Oliver Karras Eduard C. Groen

This overview shows the classification results of approaches that use the machine learning algorithms Naïve Bayes, Support Vector Machines, and Decision Trees C4.5 in combination with the machine learning features Bag of Words or Term Frequency - Inverse Document Frequency to classify user feedback as feature request.

DOI: <https://doi.org/10.48366/r112387>

Properties	Software Feature Request Detection in Issue Tracking Systems <i>User Feedback Classification - 2016</i>	Mining User Requirements from Application Store Reviews Using Frame Semantics <i>User Feedback Classification - 2017</i>	Mining Twitter Feeds for Software User Requirements <i>User Feedback Classification - 2017</i>	Automatic Classification of Non-Functional Requirements from Augmented App User Reviews <i>User Feedback Classification - 2017</i>	Bug reports simply classified <i>User Feedback Classification - 2017</i>
has dataset	https://zenodo.org/record/56907#.YKT_NudCRPY	https://mast.informatik.uni-hamburg.de/wp-content/uploads/2014/03/REJ_data.zip https://sites.google.com/site/appsuserreviews/ seel.cse.lsu.edu/data/refsq17.zip	seel.cse.lsu.edu/data/re17.zip	Not available	https://mast.informatik.uni-hamburg.de/wp-content/uploads/2014/03/REJ_data.zip
https://doi.org/10.48366/r112387					

Fig. 2: Excerpt from our comparison for Case II [35].

77 papers from the SLR by Khan et al. [18], describing the relation of the papers to five phases of RE and the CrowdRE utilities applied [36]. We are still in the process of adding the contributions from the remaining 50 papers, which is more time-consuming than for the quantitative data from Case I because of the expert judgments needed for classifying the papers' contributions. The comparison of the 27 papers makes it easy to identify, for example, the four papers that address the runtime purpose of monitoring for requirements evolution.

With the created comparisons [35], [36], we achieved our goal of acquiring and curating the detailed results of both SLRs. The knowledge-based representation in the form of comparisons has several advantages over a purely document-based representation. The comparisons are interactive and allow filtering of views by different scholarly knowledge contained in each row, even by specific value ranges of qualitative and quantitative content. The ORKG also provides a service for generating several graphical visualizations based on data in the comparisons, helping the reader understand information faster than through the large comparison table. The most important feature of the ORKG is that the added contributions and created comparisons are available to anyone. In this way, every crowd member can use the curated scholarly knowledge and created comparisons as a basis for new comparisons. Moreover, the existing comparisons can be expanded with additional scholarly knowledge from papers already included, and even with new contributions from papers added later to the ORKG. We already expanded several contributions, e.g., the results of other classifications reported in Dhinakaran et al.'s paper [40]¹⁹. For Case I, we added the details of the three crowd properties *scale*, *level of knowledge, skills & expertise*, and *roles*, which are only briefly and superficially described in the SLR [18]. For Case II, we added links to the datasets used and performance values to classification categories other than "Feature Request". This expansion is relevant to enable long-term curation. For

example, a development succeeding the SLR by Santos et al. [17] are reports of Deep Learning algorithms showing promising results in classifying user feedback [41], [42], which should be successively added to the comparison.

Despite all these advantages, the ORKG also has limitations. Most of the limitations we experienced can be attributed to the development status of the platform, which is currently in beta. Further development of the ORKG must improve interactions for the expert crowd by enabling better workflows for entering data and creating visualizations. Nevertheless, we also experienced that the project team has always responded directly to our reported issues, which we could see getting added to the GitLab issue tracker¹¹ and addressed shortly thereafter.

V. DISCUSSION

The ORKG aroused our interest as a crowdsourcing platform for applying and communicating CrowdRE research. In this experience report, we explored whether the ORKG can promote the potential of CrowdRE in open source and open research settings, taking two perspectives: that of CrowdRE researchers and that of crowd members.

Our first contribution is that we provide a comprehensive overview of the ORKG's features as a crowdsourcing platform for acquiring and curating scholarly knowledge [37], mapped to the four key activities of CrowdRE. Our findings show that the ORKG is a crowdsourcing platform offering several features that can facilitate successful CrowdRE. Although the ORKG project team has not yet consciously applied CrowdRE, they already address crucial parts of the CrowdRE cycle by motivating crowd members to participate, eliciting feedback, and monitoring context & usage data, which they analyze to derive and implement the needs and requirements of the crowd.

To motivate crowd members, the project team uses established mechanisms and incentives to boost intrinsic and extrinsic motivation (see Finding 1). Feedback is elicited

¹⁹<https://www.orkg.org/orkg/paper/R76818/R76825>

¹¹<https://gitlab.com/TIBHannover/orkg/boards/-/issues/634>

[9] Karras et al.: *Researcher or Crowd Member? Why not both! The Open Research Knowledge Graph for Applying and Communicating CrowdRE Research*. 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW), DOI: [10.1109/REW53955.2021.00056](https://doi.org/10.1109/REW53955.2021.00056), 2021.

Example: ORKG Comparison of Literature – Simulation Parameters^[10]

Comparison of Studies on Germany's Energy Supply in 2050 ★ 🔍

📅 November 2021
👤 Felix Kullmann
👤 Jan Göpfert
👤 Oliver Karras
👤 Patrick Kuckertz
👤 Sören Auer
👤 Markus Stocker
👤 Peter Markewitz

👤 Leander Kotzur
👤 Detlef Stolten

This comparison compiles the results from various studies analyzing a future low-carbon energy system for Germany. The focus of this study comparison is electricity generation. In the future, however, other essential characteristics of the respective energy system designs in the individual studies will be listed. Installed capacity is given in GW and electricity generation is given in TWh. The authors would like to thank the German Federal Government, the German State Governments, and the Joint Science Conference (GWK) for their funding and support as part of the NFDI4ing consortium. Funded by the German Research Foundation (DFG) - project number: 442146713. This work was also supported by the Helmholtz Association under the program "Energy System Design".

DOI: <https://doi.org/10.48366/r153801>

Properties	Klimaneutrales Deutschland <i>Contribution - 2020</i>	Wasserstoff-Roadmap Nordrhein-Westfalen <i>Contribution - 2020</i>	Wege zu einem klimaneutralen Energiesystem <i>Contribution - 2020</i>	Wege für die Energiewende <i>Contribution - 2019</i>	Den Weg der Energiewende <i>Contribution - 2019</i>
has energy_sources	all sources	all sources	all sources	all sources	
	bioenergy	bioenergy	bioenergy	bioenergy	
	geothermics	geothermics	geothermics	geothermics	
	hydropower	hydropower	hydropower	hydropower	
	import	import			
	net import	net import			

<https://doi.org/10.48366/r153801>

O. Karras et al. | Organizing Scientific Knowledge From Energy System Research Using the ORKG

and GHG scenarios [14] (see Figure 1). In contrast to the traditional way of publishing an overview of scenarios within a publication, ORKG *comparisons* provide the benefit that they are versionable and can thus be continuously (re)used, updated, and expanded. When researchers publish new scenarios as factsheets or in publications, the ORKG *comparisons* can be easily extended by describing the new scenarios using the same ORKG *templates*, adding the new ORKG *contribution* to the respective ORKG *comparison*, and publishing the updated ORKG *comparison* as a new version. The ORKG also supports the supplementation of ORKG *comparisons* by creating visualizations based on the data contained therein either directly from the web frontend or via various access points, such as a REST API, a Python or R package, or a SPARQL endpoint, for example in combination with a Jupyter notebook.

Figure 1. ORKG comparison of 25 scenarios from GHG studies for Germany [14].

In addition, we established an ORKG *observatory* on Energy System Research¹². The ORKG *observatory* serves as a central access point to all related curated publications, comparisons, and visualizations so that other researchers can easily explore the content. For example, Auer et al. [15] already reused the curated scientific knowledge from our two ORKG *comparisons* by identifying and answering further natural language competency questions from domain experts beyond the previous consideration. For this purpose, they specified the competency question as SPARQL query (see Listing 1). We executed this query on the SPARQL endpoint and visualized the results in Figure 2. In particular, these results show that average energy supply from photovoltaics and onshore wind power increased approximately fourfold from the 2006 – 2010 interval to the 2016 – 2020 interval.

¹²https://orkg.org/observatory/Energy_System_Research

[10] Karras et al.: *Organizing Scientific Knowledge From Energy System Research Using the Open Research Knowledge Graph*. 1st NFDI4Energy Conference, DOI: [10.48550/arXiv.2401.13365](https://doi.org/10.48550/arXiv.2401.13365), 2024.

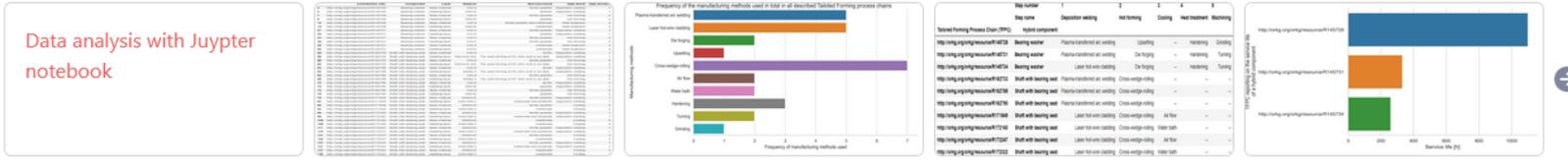
Example: ORKG Comparison of Literature – Process Variants^[11]

Tailored Forming Process Chain for the Manufacturing of Hybrid Components with Bearing Raceways Using Different Material Combinations ★

May 2022 Oliver Karras Laura Budde Paulina Merkel

This comparison provides an overview of Tailored Forming process chains carried out in the CRC 1153 "Tailored Forming" for the manufacturing of hybrid components with bearing raceways using different material combinations. A variety of materials combinations is used for the Tailored Forming of such hybrid components. The comparison shows the combined materials, the entire Tailored Forming process chain with its individual steps, and the resulting hybrid component with its qualities. For each step, the comparison shows the measurement methods performed and their results for the individual qualities of the hybrid components. In this way, the comparison shows how different qualities of the hybrid components change during the process due to the manufacturing methods used.

DOI: <https://doi.org/10.48366/r187049>



Properties	Investigation of the material combination 20MnCr5 and X45CrSi9-3 in the Tailored Forming of shafts with bearing seats <i>Air-cooled 2 Cladding Layer Component (x45CrSi9-3) - 2022</i>	Investigation of the material combination 20MnCr5 and X45CrSi9-3 in the Tailored Forming of shafts with bearing seats <i>Air-cooled 3 Cladding Layer Component (x45CrSi9-3) - 2022</i>	Investigation of the material combination 20MnCr5 and X45CrSi9-3 in the Tailored Forming of shafts with bearing seats <i>Water-cooled 2 Cladding Layer Component (x45CrSi9-3) - 2022</i>	Investigation of the material combination 20MnCr5 and X45CrSi9-3 in the Tailored Forming of shafts with bearing seats <i>Water-cooled 3 Cladding Layer Component (x45CrSi9-3) - 2022</i>	Cross-welded rolling material <i>1 Claddi (100Cr6)</i>
has material/material					
↳ belongs to material group*	Steel	Steel	Steel	Steel	
↳ has type*	20MnCr5 X45CrSi9-3	20MnCr5 X45CrSi9-3			

<https://doi.org/10.48366/r187049>

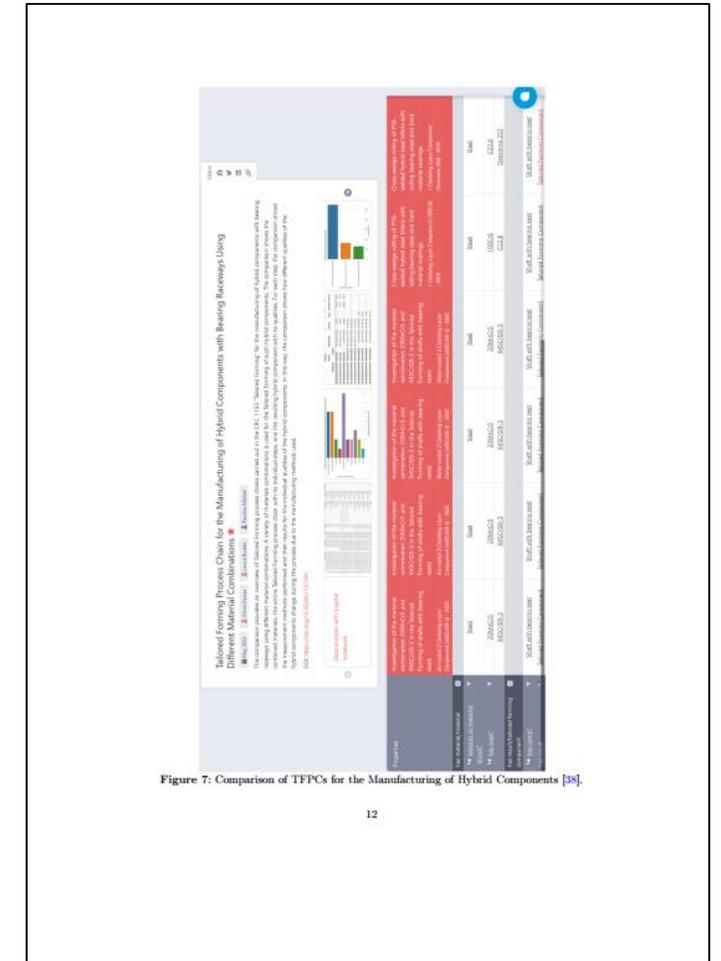


Figure 7: Comparison of TFPCs for the Manufacturing of Hybrid Components [8].

[11] Karras et al.: *Organizing Scientific Knowledge from Engineering Sciences Using the Open Research Knowledge Graph: The Tailored Forming Process Chain Use Case. Under Review* at Data Science Journal, 2024.

Example: ORKG Comparison of Literature: Software Features^[12]

Comparison of Hyperparameter Optimization Tools ★

January 2023 Oliver Karras Difan Deng Marius Lindauer

This comparison looks at the capabilities of various optimization tools for hyperparameter optimization, focusing on the following: Complex hyperparameter space, multi-objective, multi-fidelity, instances, command-line interface, and parallel computing. In addition, this comparison provides further information about the respective optimization tool, such as the locations of the code repository, README, software documentation and issue tracker, license and programming language.

DOI: <https://doi.org/10.48366/r281265>

Properties	Tool	BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization <i>BoTorch - 2020</i>	OpenBox: A Generalized Black-box Optimization Service <i>OpenBox - 2021</i>	BOHB: Robust and Efficient Hyperparameter Optimization at Scale <i>HpBandSter - 2018</i>	SMAC3: A Versatile Bayesian Optimization Package for Hyperparameter Optimization <i>SMAC3 - 2021</i>
result/software					
name*		BoTorch	OpenBox	HpBandSter	SMAC3
software features*	Complex Hyperparameter Space	Multi-Fidelity	Complex Hyperparameter Space	Complex Hyperparameter Space	command-line interface
	Multi-Objective	Multi-Objective	Multi-Objective	Multi-Fidelity	Complex Hyperparameter Space
	parallel computing	parallel computing	parallel computing	parallel computing	Instances
					Multi-Fidelity
					Multi-Objective
					parallel computing
complex hyperparameter space*		✗	✓	✓	✓
multi-objective*		✓	✓	✗	✓
multi-fidelity*		✓	✗		
instance*		✗	✗		

<https://doi.org/10.48366/r281265>

Command-Line Interface

SMAC can not only be executed within a python file but also from the commandline. Consequently, not only algorithms in python can be optimized, but implementations in other languages as well.

Note

Command-line interface has been temporarily disabled in v2.0. Please fall back to v1.4 if you need it.

Comparison

The following table provides an overview of SMAC's capabilities in comparison with other optimization tools.

Package	Complex Hyperparameter Space	Multi-Objective	Multi-Fidelity	Instances	Command-Line Interface
HyperMapper	✓	✓	✗	✗	✗
Optuna	✓	✓	✓	✗	✓
Hyperopt	✓	✗	✗	✗	✓
BoTorch	✗	✓	✓	✗	✗
OpenBox	✓	✓	✗	✗	✗
HpBandSter	✓	✗	✓	✗	✗
SMAC	✓	✓	✓	✓	✓

[12] Lindauer et al.: SMAC3 Documentation, URL: https://automl.github.io/SMAC3/v1.4.0/getting_started/package_overview.html#comparison, 2024.

Goal: ORKG Comparison on Reported Empirical Research

Comparison | 4 contributions

Overview of Reported Empirical Research in Requirements Engineering Publications from 2021

This comparison shows an overview of reported empirical research in publications of the IEEE International Requirements Engineering Conference from the year 2021

Properties	On the impact of using different templates on creating and understanding user stories Contribution 1 - 2021	Ambiguity and Generality in Natural Language Privacy Policies Contribution 1 - 2021	Environment-Driven Abstraction Identification for Requirements-Based Testing Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research question/research question				
↳ question*	Do different user story templates have an impact on the creation of user stories? Do different user story templates have an impact on the understanding of user stories?	To what extent does the classification model reduce the manual ontology construction effort? What is the effect of missing transitive hypenymy on classification performance?	No question.	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
↳ highlighted in text*	✓ ✓	✓ ✓	✗	✓ ✓ ✓
↳ hidden in text*	✗ ✗	✗ ✗	✗	✗ ✗ ✗
data collection method/data collection method				
↳ method*	experiment	experiment	experiment	experiment
↳ number of participants*	41	0	0	30
data analysis method/data analysis method/method*	descriptive statistics inferential statistics	descriptive statistics machine learning	comparative analysis descriptive statistics machine learning	descriptive statistics inferential statistics thematic analysis
threats to validity/threats to validity				
↳ conclusion validity*	✓	✗	✗	✗
↳ construct validity*	✓	✓	✓	✓
↳ external validity*	✓			
↳ internal validity*	✓			

<https://sandbox.orkg.org/comparison/R369109>

Topic: Reported Empirical Research

- Empirical Research
 - *Research question*
 - Question
 - Hidden in text
 - Highlighted in text
 - *Data Collection Method*
 - Method
 - Number of participants
 - *Data Analysis Method*
 - Method
 - *Threats to Validity*
 - Construct Validity
 - Internal Validity
 - External Validity
 - Conclusion Validity

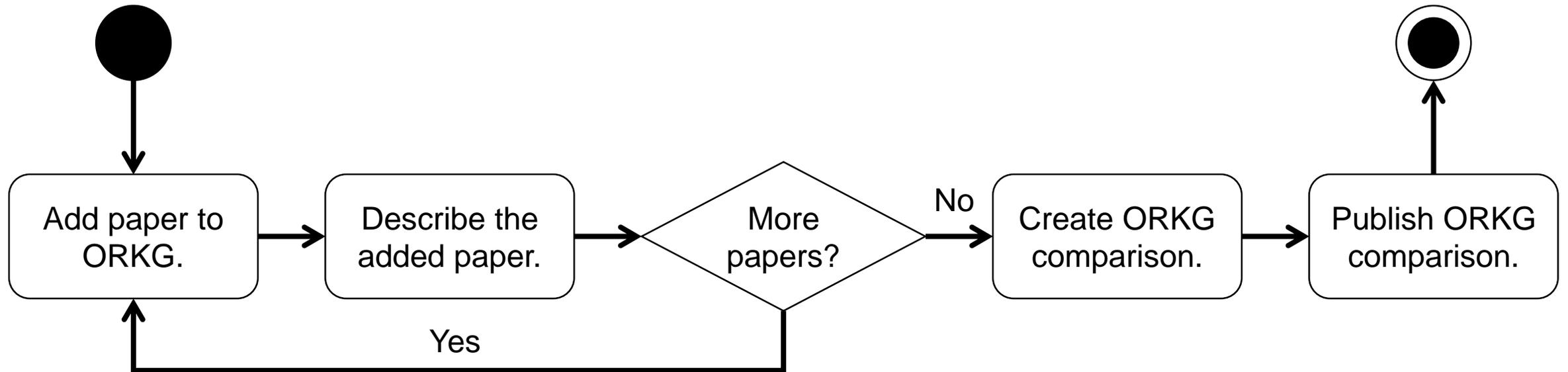
The screenshot displays the ORKG interface for a specific paper. At the top, the ORKG logo and navigation menu are visible. The paper title is "Environment-Driven Abstraction Identification for Requirements-Based Testing". Below the title, there are metadata tags for "September 2021", "5 citations", "Software Engineering", and several authors. The publication information is "Published in: 2021 IEEE 29th International Requirements Engineering Conference (RE)" and the DOI is "https://doi.org/10.1109/re51729.2021.00029".

The main content area shows a "Contribution 1" section with a table of applied templates. The table lists various research-related terms and their corresponding values or links:

Applied template	Value
data analysis method	comparative analysis machine learning descriptive statistics
data collection method	experiment
research question	No question.
threats to validity	Construct validity: true, Internal validity: true, External validity: true Conclusion validity: false

On the right side of the interface, there are additional options such as "Add to comparison", "Provenance", "Timeline", "Added on" (08 May 2024), "Added by" (Oliver Karras), "Contributors" (Oliver Karras), and "Assign to observatory".

Overall Process for Creating an ORKG Comparison



Remark:

Instead of going through the loop, we will work **collaboratively** in the tutorial as conceived by ORKG. We will take all papers from all participants and compare them with each other in an ORKG comparison.

An ORKG comparison requires **at least two** publications.

1. Open ORKG Website

The screenshot shows the ORKG website interface. At the top, there is a navigation bar with the ORKG logo, menu items (View, Tools, About), a dropdown menu for 'NFDI4DataScience', a search bar, and buttons for '+ Add new' and 'Sign in'. Below the navigation bar, the main heading reads 'Scholarly Knowledge. Structured.' followed by a sub-heading: 'The Open Research Knowledge Graph (ORKG) aims to describe research papers in a structured manner. With the ORKG, papers are easier to find and compare. Play video'. A section titled 'Browse by research field' contains five red buttons: 'Arts and Humanities' (443 papers - 34 comparisons), 'Engineering' (3434 papers - 356 comparisons), 'Life Sciences' (4169 papers - 206 comparisons), 'Physical Sciences & Mathematics' (15830 papers - 745 comparisons), and 'Social and Behavioral Sciences' (944 papers - 174 comparisons). Below this, there are tabs for 'Comparisons', 'Papers', 'Visualizations', 'Reviews', and 'Lists'. The 'Comparisons' tab is active, showing a list of comparisons, including one titled 'Overview on studies about the role of (social) media in political polarization across the years 2011 to 2020'. On the right side, there is a 'Latest Mastodon posts' section with a post from 'Open Research Knowledge Graph' (@orkg@mastodon.social) with the text 'Wonder why #reborn papers come with a butterfly 🦋?'.

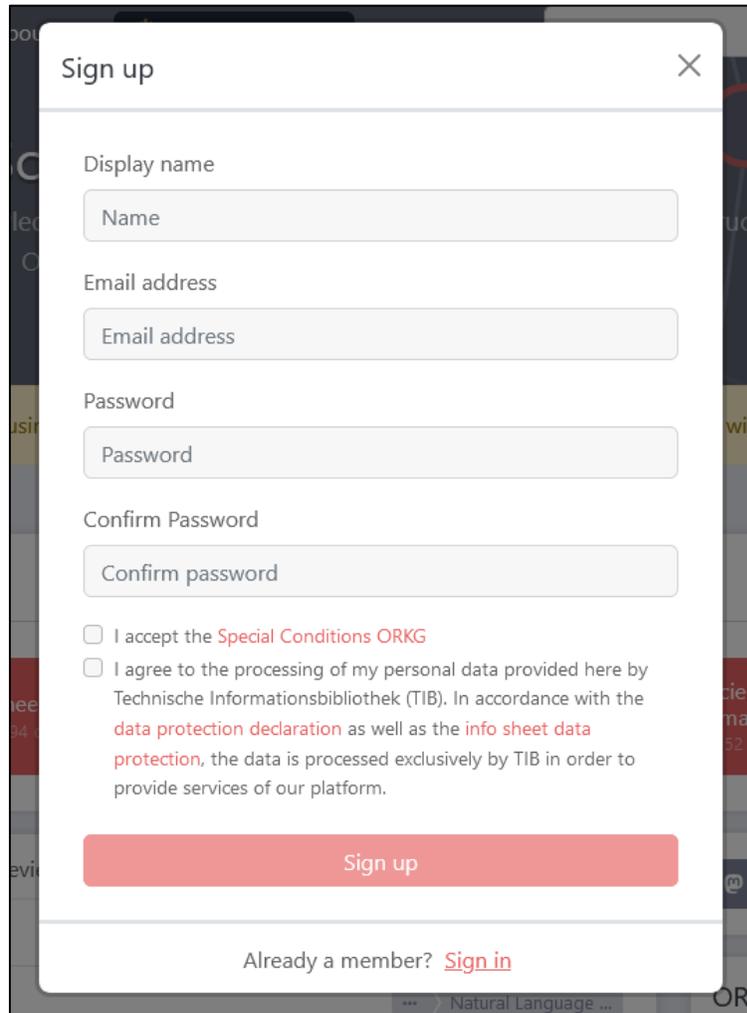
Remark:

For the tutorial, we use the **ORKG test environment** (<https://sandbox.orkg.org/>) to try everything without hesitation. You can find the official ORKG at <https://orkg.org/>.

Get your exemplary paper: <https://bit.ly/3UGFwhj>

Use <https://sandbox.orkg.org/>!
NOT <https://orkg.org/>!

2. Sign Up & Sign In



Sign up

Display name

Name

Email address

Email address

Password

Password

Confirm Password

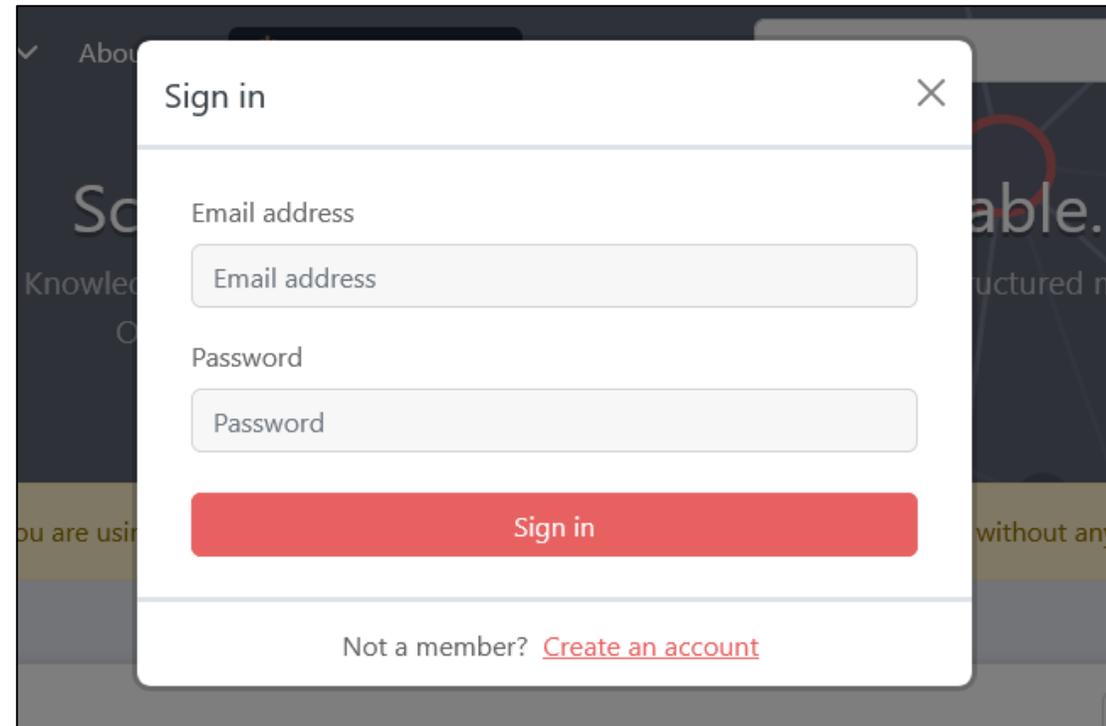
Confirm password

I accept the [Special Conditions ORKG](#)

I agree to the processing of my personal data provided here by Technische Informationsbibliothek (TIB). In accordance with the [data protection declaration](#) as well as the [info sheet data protection](#), the data is processed exclusively by TIB in order to provide services of our platform.

Sign up

Already a member? [Sign in](#)



Sign in

Email address

Email address

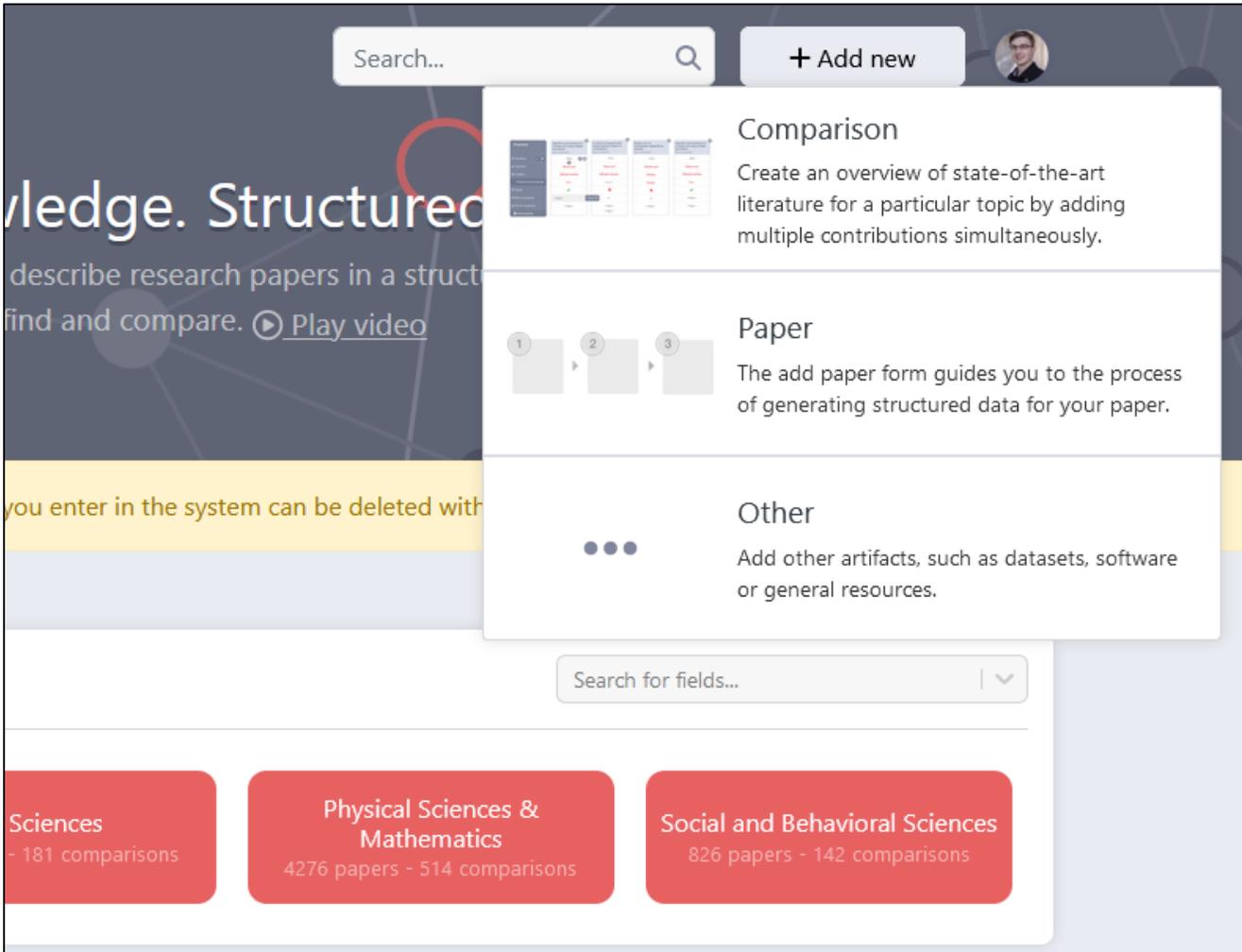
Password

Password

Sign in

Not a member? [Create an account](#)

3. Add a New Comparison to the ORKG



The screenshot shows the ORKG interface. At the top, there is a search bar with the text "Search..." and a magnifying glass icon, followed by a "+ Add new" button and a user profile picture. A dropdown menu is open, showing three options: "Comparison", "Paper", and "Other". The "Comparison" option is highlighted with a red circle. Below the dropdown, there is a "Search for fields..." dropdown menu. At the bottom, there are three red buttons representing different scientific fields: "Sciences" (181 comparisons), "Physical Sciences & Mathematics" (4276 papers - 514 comparisons), and "Social and Behavioral Sciences" (826 papers - 142 comparisons).

1. Click on “+Add new”
2. Select “Comparison”

Remark:

We could also start by adding individual papers, but since we want to create a comparison at the end, it is easy to start immediately.

4. Go to the Contribution Editor

ORKG View Tools About NFDI4DataScience Search... + Add new

Add comparison

How to make an ORKG comparison - An Example from Virology

Comparisons in ORKG provide an overview of state-of-the-art literature for a particular topic. Comparisons are dynamic and FAIR. A comparison is created from contributions, [view example of comparison](#). To create your own comparisons in ORKG, you can either import existing data (via CSV import) or start from scratch by adding your own contributions. This page guides you in creating new comparisons.

- 1. Existing data**
In case you have existing data, you can import this via the CSV import tool. This is especially helpful if you already have a large file in which related work is compared.
[Go to CSV import tool](#)
- 2. Contribution editor**
If you don't have existing data, go to the contribution editor to add contributions that will be used in the comparison. After creating contributions, you can create a comparisons.
[Go to contribution editor](#)
- 3. Publish comparison**
Once you are done editing contributions, you can create and publish a comparison. Published comparisons are persistent so they are perfectly suitable for publications.
[Publish comparison](#)

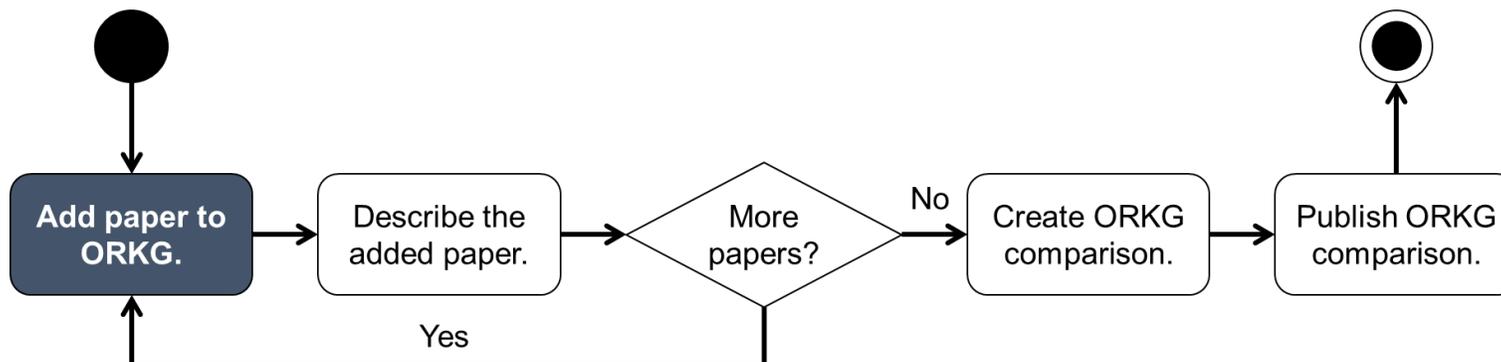
1. Click on “Go to contribution editor”

Remark:

When the data is already extracted, e.g., in a spreadsheet, we could use CSV import tool. However, the tool is a prototype and still requires a lot of manual work to map the data.

5. Add a Paper

The screenshot shows the ORKG Contribution editor interface. At the top left is the ORKG logo. Navigation links include 'View', 'Tools', and 'About'. A dropdown menu is set to 'NFDI4DataScience'. A search bar is on the right, followed by a red '+ Add new' button and a user profile picture. A yellow warning banner reads: 'Warning: You are using a testing environment. Data you enter in the system can be deleted without any notice.' Below this is the 'Contribution editor' section with 'View comparison' and '+ Add contribution' buttons. A light blue instruction box says: 'Start adding contributions by clicking the button *Add contribution* on the right'.



1. Click on “Add contribution”

5. Add a Paper: Enter DOI (or Paper Title)

Add contribution

Paper title or DOI

Search contributions by paper title or DOI...

Select the contributions you want to add or you click on + if you want to create a new contribution for an existing paper

- From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation +
- Environment-Driven Abstraction Identification for Requirements-Based Testing +
- Ambiguity and Generality in Natural Language Privacy Policies +
- On the impact of using different templates on creating and understanding user stories +
- testadas +
 - Contribution 1
 - Contribution 2

+ Load more

Add contributions

Add contribution

Paper title or DOI

10.1109/RE51729.2021.00009

There are no results, please try a different search term or [Add new paper](#)

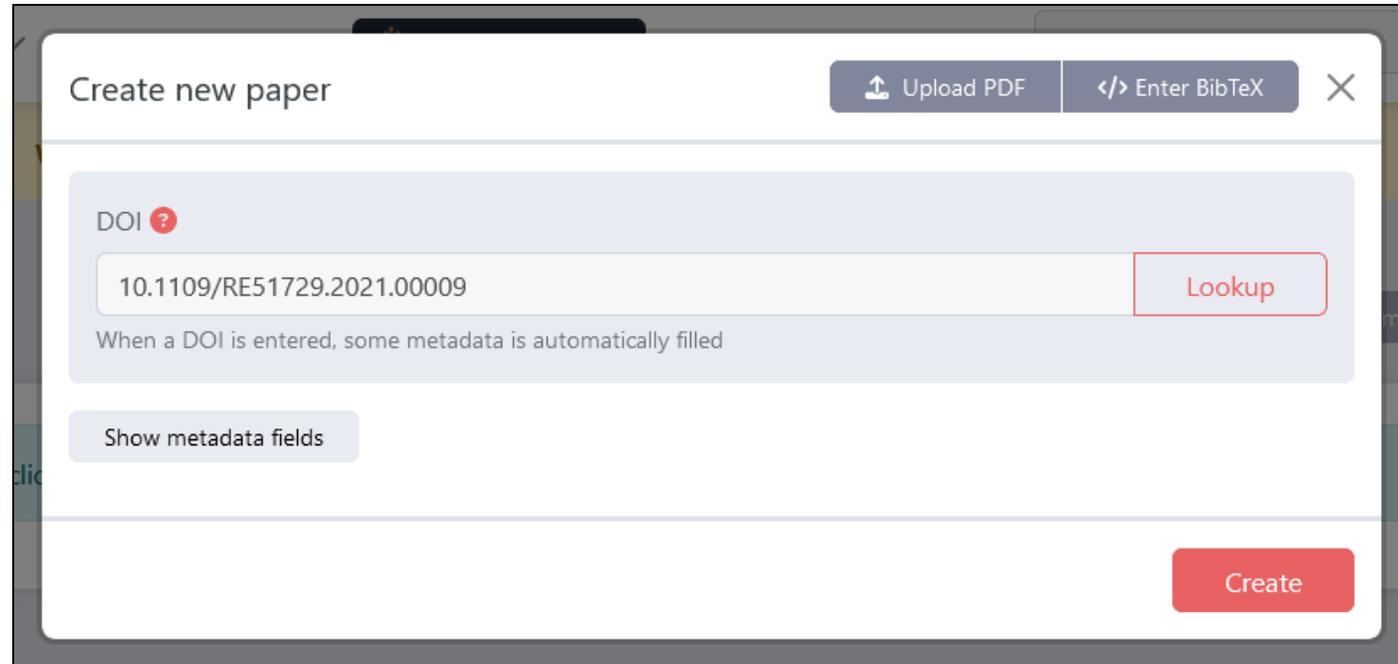
Add contributions

1. Click on “Add new paper”

Remark:

If the paper is already in the ORKG, it is shown and we can either reuse an existing contribution or create a new one.

5. Add a Paper: Lookup Paper by DOI

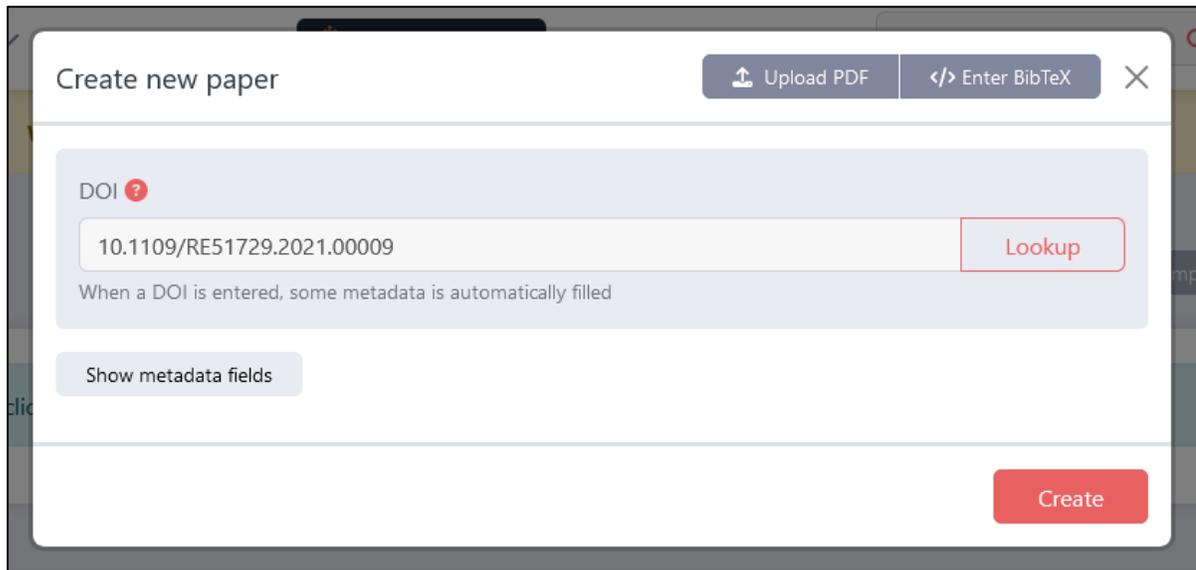


The screenshot shows a web interface for creating a new paper. At the top, there is a title "Create new paper" and two buttons: "Upload PDF" and "Enter BibTeX". Below this is a form with a "DOI" label and a red question mark icon. A text input field contains the DOI "10.1109/RE51729.2021.00009". To the right of the input field is a red "Lookup" button. Below the input field, there is a note: "When a DOI is entered, some metadata is automatically filled". Below the note is a "Show metadata fields" button. At the bottom right of the form is a red "Create" button.

Remark:

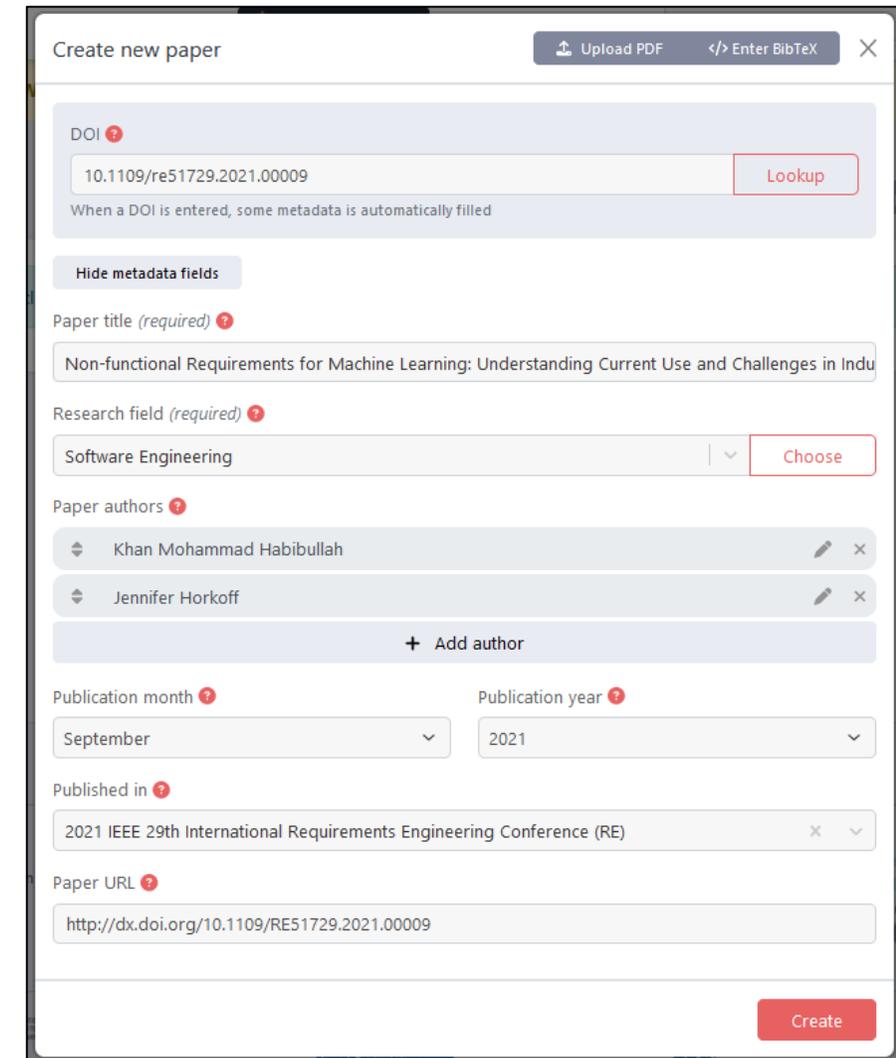
ORKG attempts to retrieve the metadata automatically. If the retrieval is not successful, you can also manually enter or correct the metadata.

5. Add a Paper: Lookup Paper by DOI



The screenshot shows the 'Create new paper' form with the following elements:

- Buttons: 'Upload PDF', 'Enter BibTeX', and 'Create'.
- Field: 'DOI' with a red question mark icon.
- Input: '10.1109/RE51729.2021.00009'.
- Button: 'Lookup'.
- Text: 'When a DOI is entered, some metadata is automatically filled'.
- Button: 'Show metadata fields'.



The screenshot shows the 'Create new paper' form with the following elements:

- Buttons: 'Upload PDF', 'Enter BibTeX', and 'Create'.
- Field: 'DOI' with a red question mark icon.
- Input: '10.1109/re51729.2021.00009'.
- Button: 'Lookup'.
- Text: 'When a DOI is entered, some metadata is automatically filled'.
- Button: 'Hide metadata fields'.
- Field: 'Paper title (required)' with a red question mark icon.
- Input: 'Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Indu'.
- Field: 'Research field (required)' with a red question mark icon.
- Input: 'Software Engineering'.
- Button: 'Choose'.
- Field: 'Paper authors' with a red question mark icon.
- Input: 'Khan Mohammad Habibullah'.
- Input: 'Jennifer Horkoff'.
- Button: '+ Add author'.
- Field: 'Publication month' with a red question mark icon.
- Input: 'September'.
- Field: 'Publication year' with a red question mark icon.
- Input: '2021'.
- Field: 'Published in' with a red question mark icon.
- Input: '2021 IEEE 29th International Requirements Engineering Conference (RE)'.
- Field: 'Paper URL' with a red question mark icon.
- Input: 'http://dx.doi.org/10.1109/RE51729.2021.00009'.

1. Click on “Lookup”
2. Check metadata fetched
3. Add a “Research field”. We use “Software Engineering”.
4. Click on “Create”

6. Describe a Paper

Contribution editor

View comparison | Add contribution

Properties ✓ Saved Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry ✕

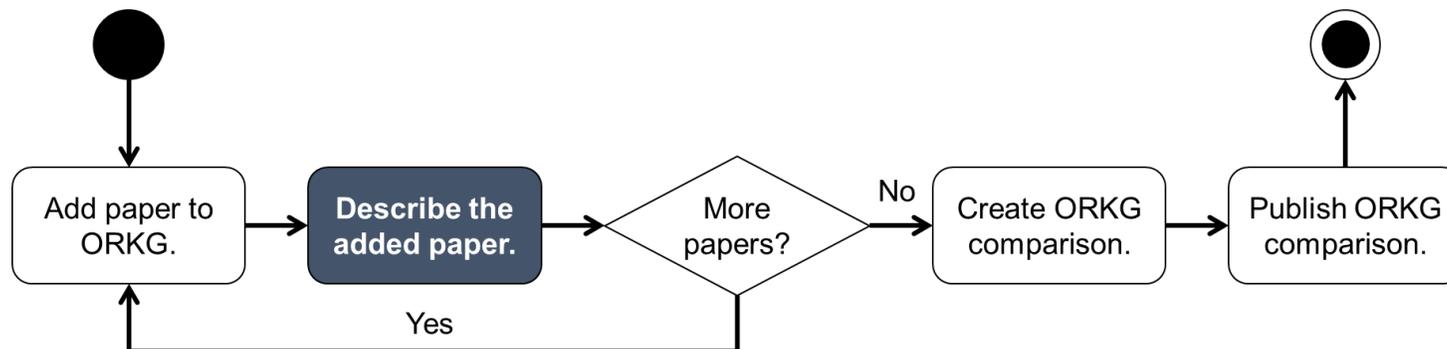
Contribution 1

Start adding properties or use templates by using the buttons below

+ Add property | Templates

Suggested properties

+ research problem Problem



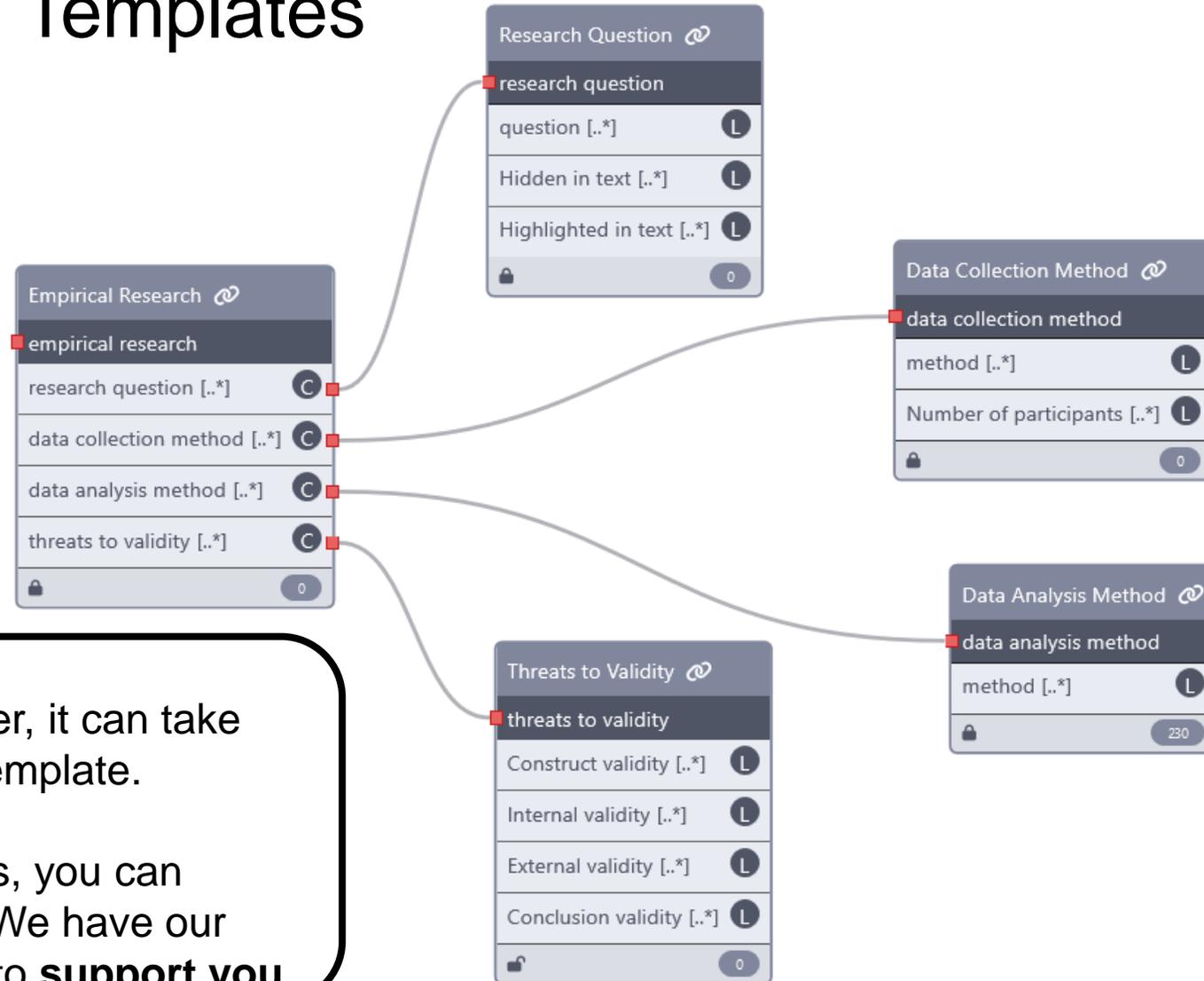
Remark:

Instead of adding all properties individually, we use an ORKG template.

6. Describe a Paper: ORKG Templates

An **ORKG template** specifies the **structure** and **terminology** used to describe a **publication** in the ORKG. Its use **ensures FAIR data** that is **consistent** and **comparable** across publications.

<https://sandbox.orkg.org/template/R369028>

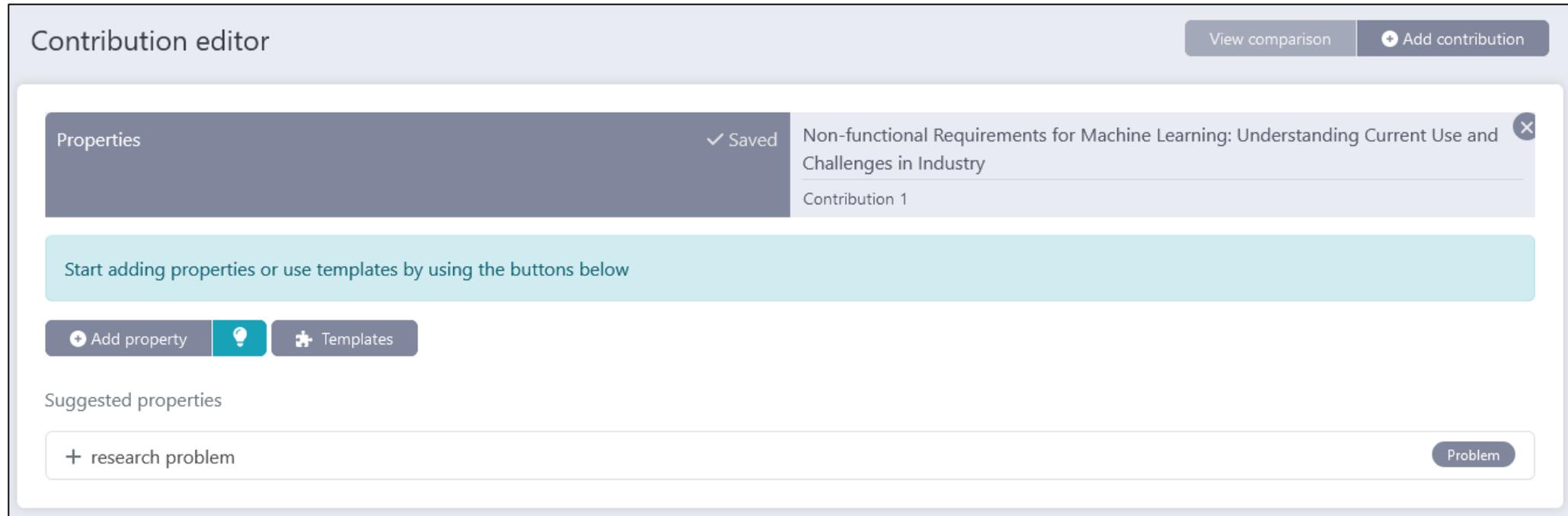


Remark:

Every user can create ORKG templates. However, it can take some time and practice to build a good ORKG template.

If you want to use the feature and have problems, you can always contact the ORKG team, especially me. We have our own **Curation and Community Building team** to support you.

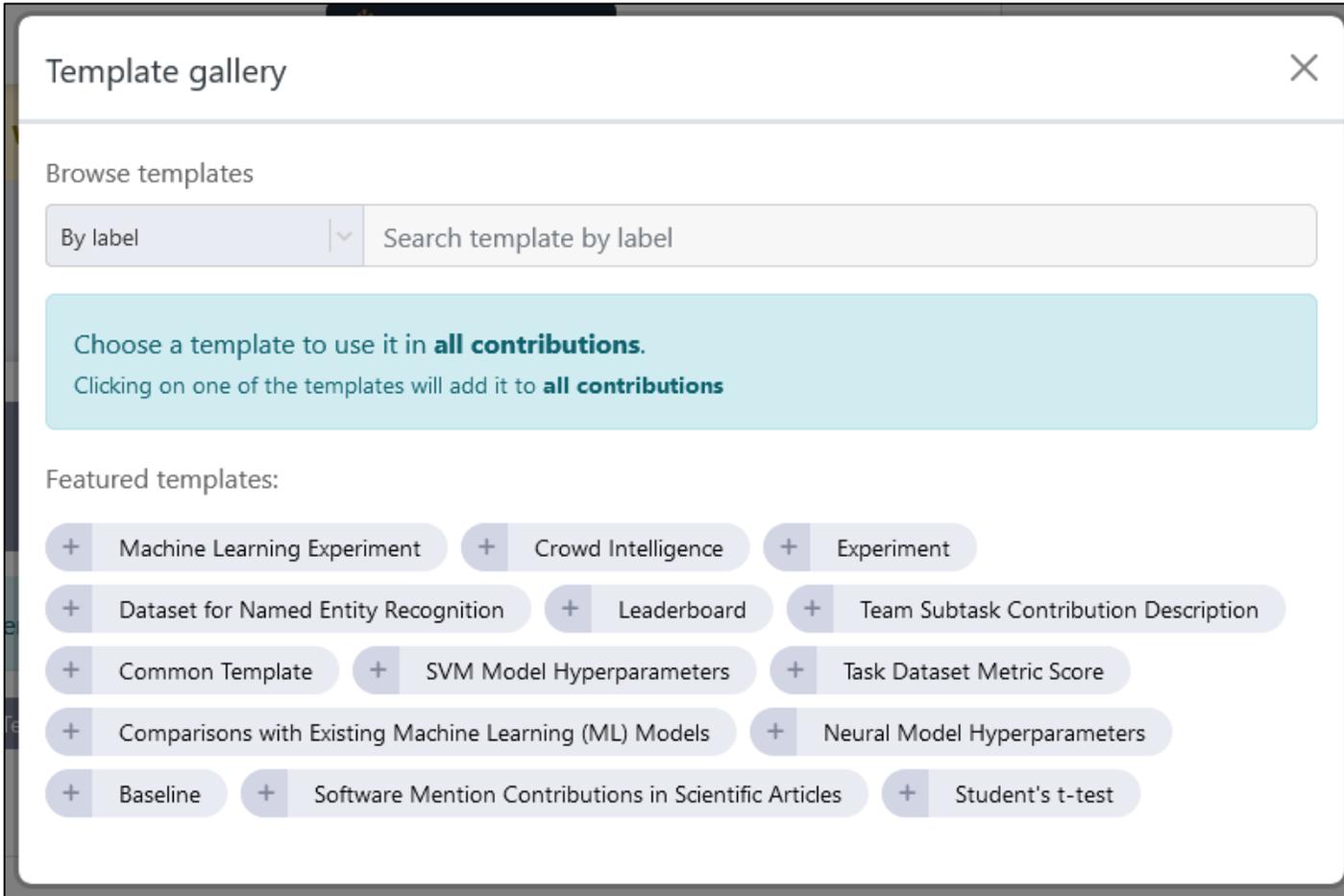
6. Describe a Paper: Add Template “Empirical Research”



The screenshot shows the 'Contribution editor' interface. At the top right, there are two buttons: 'View comparison' and '+ Add contribution'. Below this, a dark grey bar displays 'Properties' on the left, a '✓ Saved' status in the middle, and the title 'Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry' on the right, with a close icon (X) at the far right. Underneath the title, it says 'Contribution 1'. A light blue banner below the title bar contains the text 'Start adding properties or use templates by using the buttons below'. Below the banner are three buttons: '+ Add property' (with a plus icon), a light blue button with a lightbulb icon, and '+ Templates' (with a puzzle piece icon). Under the 'Suggested properties' section, there is a search bar containing '+ research problem' and a 'Problem' button on the right.

1. Click on “Templates”

6. Describe a Paper: Add Template “Empirical Research”



Template gallery

Browse templates

By label Search template by label

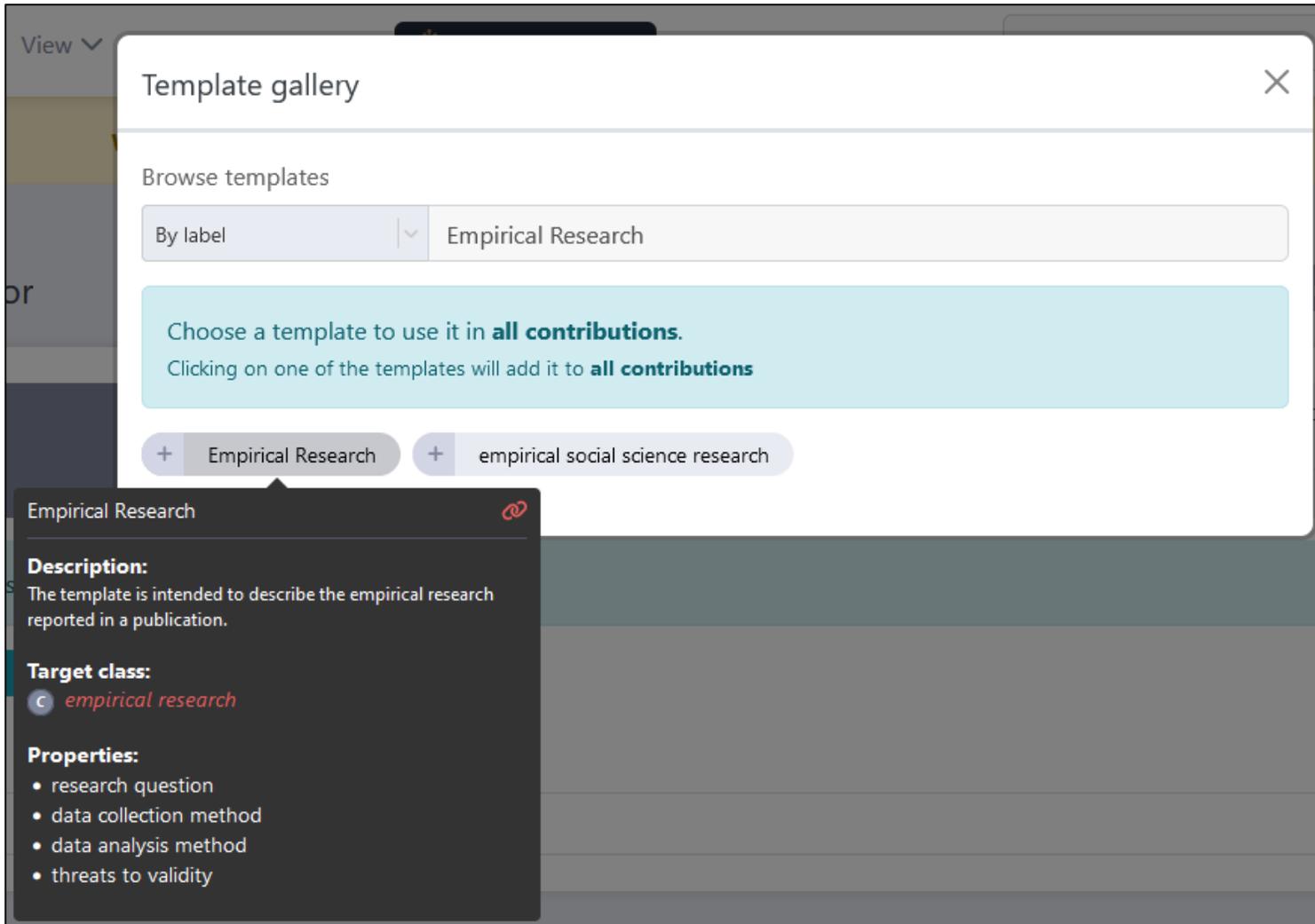
Choose a template to use it in **all contributions**.
Clicking on one of the templates will add it to **all contributions**

Featured templates:

- + Machine Learning Experiment
- + Crowd Intelligence
- + Experiment
- + Dataset for Named Entity Recognition
- + Leaderboard
- + Team Subtask Contribution Description
- + Common Template
- + SVM Model Hyperparameters
- + Task Dataset Metric Score
- + Comparisons with Existing Machine Learning (ML) Models
- + Neural Model Hyperparameters
- + Baseline
- + Software Mention Contributions in Scientific Articles
- + Student's t-test

1. Search for “Empirical Research”

6. Describe a Paper: Add Template “Empirical Research”



The screenshot shows a 'Template gallery' window with a search bar containing 'Empirical Research'. Below the search bar, there is a light blue instruction box: 'Choose a template to use it in **all contributions**. Clicking on one of the templates will add it to **all contributions**'. At the bottom of the gallery, two template buttons are visible: '+ Empirical Research' and '+ empirical social science research'. A tooltip is open over the 'Empirical Research' button, displaying the following information:

Empirical Research 

Description:
The template is intended to describe the empirical research reported in a publication.

Target class:
 *empirical research*

Properties:

- research question
- data collection method
- data analysis method
- threats to validity

1. Select “Empirical Research”

6. Describe a Paper: Fill out the Table

The screenshot shows the 'Contribution editor' interface. At the top right, there are buttons for 'View comparison' and '+ Add contribution'. The main area is divided into two columns. The left column, titled 'Properties', contains a list of properties: 'data analysis method', 'data collection method', 'research question', and 'threats to validity'. A '✓ Saved' indicator is visible in the top right of this column. The right column, titled 'Contribution 1', contains a table with one row. The first cell of this row is empty and has a tooltip that says 'Add value' with a '+' icon below it. Below the table, there are buttons for '+ Add property', a lightbulb icon, and '+ Templates'. At the bottom, there is a section for 'Suggested properties' with a button for '+ research problem' and a 'Problem' label.

1. Click “Add value” for adding the data analysis method

6. Describe a Paper: Enter the Data Analysis Method(s)

The screenshot shows the ORKG Contribution editor interface. A modal window titled "View existing resource: Data Analysis Method" is open, displaying the "Applied template: Data Analysis Method". Below the template, there is a text input field containing the word "method" and a plus sign (+) button. The background interface includes a "Properties" sidebar with links for "data_analysis_method", "data_collection_method", "research_question", and "threats_to_validity". At the bottom, there are buttons for "Add property", "Templates", and "Suggested properties" with a "research problem" option.

1. Click the +
2. Enter method

6. Describe a Paper: Enter the Data Analysis Method(s)

The screenshot shows the ORKG Contribution editor interface. A modal dialog box titled "View existing resource: Data Analysis Method" is open. The dialog has a close button (X) and an "Open resource" link. Below the title, there are buttons for "Templates", "Help", and "Preferences". A section labeled "Applied template: Data Analysis Method" is visible. Below this, there is a text input field with the label "method" and the value "thematic analysis". To the right of the input field are buttons for "Cancel" and "Create". At the bottom of the dialog is a "+ Add property" button. The background shows the "Contribution editor" with a "Properties" panel on the left containing links for "data analysis method", "data collection method", "research question", and "threats to validity". A "Suggested properties" section at the bottom left shows a "+ research problem" button. The top right of the interface has a search icon, a "+ Add new" button, and a user profile icon.

1. Click “Create”
2. Close dialog

6. Describe a Paper: Enter the Data Analysis Method(s)

The screenshot shows the ORKG Contribution editor interface. At the top, there is a navigation bar with the ORKG logo, menu items (View, Tools, About), a user profile dropdown (NFDI4DataScience), a search bar, and a '+ Add new' button. A yellow warning banner states: "Warning: You are using a testing environment. Data you enter in the system can be deleted without any notice." Below this is the 'Contribution editor' section, which includes a 'View comparison' button and a '+ Add contribution' button. The main editing area is divided into two columns. The left column, titled 'Properties', contains a list of property labels: 'data analysis method', 'data collection method', 'research question', and 'threats to validity'. The right column shows the contribution title 'Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry' and the label 'Contribution 1'. The 'data analysis method' field is highlighted in red and contains the text 'thematic analysis'. At the bottom left, there are buttons for '+ Add property', a lightbulb icon, and '+ Templates'. At the bottom right, there is a 'Suggested properties' section with a search bar containing '+ research problem' and a 'Problem' button.

If you have more analysis methods, **repeat** the process.

Otherwise, continue with the next property.

6. Describe a Paper: Enter the Data Collection Method(s)

The screenshot shows the ORKG Contribution editor interface. A modal window titled "View existing resource: Data Collection Method" is open, displaying the "Data Collection Method" template. The modal includes a search bar, "Templates", "Help", and "Preferences" buttons. Below the search bar, the applied template is shown as "Data Collection Method". The modal contains two input fields: "method" and "Number of participants", each with a plus sign (+) button to its right. At the bottom of the modal is a "+ Add property" button. The background shows the "Contribution editor" with a "Properties" panel on the left containing links for "data analysis method", "data collection method", "research question", and "threats to validity". At the bottom of the editor, there are buttons for "+ Add property", "Templates", and "Suggested properties" with a "+ research problem" button.

1. Click the +
2. Enter method
3. Click the +
4. Enter number of participants

6. Describe a Paper: Enter the Data Collection Method(s)

The screenshot shows the ORKG Contribution editor interface. A modal dialog titled "View existing resource: Data Collection Method" is open. The dialog has a close button (X) and an "Open resource" link. It contains a "Templates" button, a "Help" button, and a "Preferences" button. Below these is a section for "Applied template: Data Collection Method". The dialog has two input fields: "method" with a dropdown set to "Text" and the value "interview", and "Number of participants" with a dropdown set to "Integer" and the value "10". Each field has a lightbulb icon, a "Cancel" button, and a "Create" button. At the bottom of the dialog is a "+ Add property" button. The background shows the "Contribution editor" with a "Properties" panel on the left containing links for "data analysis method", "data collection method", "research question", and "threats to validity". There is also a "Suggested properties" section with a "+ research problem" button.

1. Click “Create”
2. Click “Create”
3. Close dialog

6. Describe a Paper: Enter the Data Collection Method(s)

The screenshot shows the ORKG Contribution editor interface. At the top, there is a navigation bar with the ORKG logo, menu items (View, Tools, About), a user profile icon, and a search bar. A yellow warning banner states: "Warning: You are using a testing environment. Data you enter in the system can be deleted without any notice." The main area is titled "Contribution editor" and contains a table with the following content:

Properties	Value
data analysis method	thematic analysis
data collection method	interview
research question	
threats to validity	

Below the table, there are buttons for "Add property", "Templates", and "Suggested properties". The "Suggested properties" section shows a "+ research problem" button with a "Problem" label.

If you have more collection methods, **repeat** the process.

Otherwise, continue with the next property.

6. Describe a Paper: Enter the Research Question(s)

The screenshot shows the ORKG Contribution editor interface. A modal dialog box titled "View existing resource: Research Question" is open, displaying the "Applied template: Research Question". The dialog contains three input fields: "question", "Hidden in text", and "Highlighted in text", each with a plus sign (+) to its right. Below these fields is a "+ Add property" button. The background interface shows the "Contribution editor" with a "Properties" panel on the left containing links for "data analysis method", "data collection method", "research question", and "threats to validity". At the bottom, there are buttons for "+ Add property", "Templates", and "Suggested properties" with a search bar containing "+ research problem".

1. Click the +
2. Enter question
3. Click the +
4. Enter, if the question is hidden in text
5. Click the +
6. Enter, if the question is highlighted in text

6. Describe a Paper: Enter the Research Question(s)

The screenshot shows the ORKG Contribution editor interface. A modal dialog titled "View existing resource: Research Question" is open. The dialog has a header with "Open resource" and a close button. Below the header are three tabs: "Templates", "Help", and "Preferences". The main content area shows the applied template "Research Question" and three input fields:

- question**: Text input with the value "nt treatment of NFRs in ML in industry?".
- Hidden in text**: Boolean input with the value "False".
- Highlighted in text**: Boolean input with the value "True".

Each input field has a "Cancel" button and a "Create" button. At the bottom of the dialog is a "+ Add property" button. The background shows the "Contribution editor" interface with a "Properties" sidebar and a "Suggested properties" section.

1. Click “Create”
2. Click “Create”
3. Click “Create”
4. Close dialog

6. Describe a Paper: Enter the Research Question(s)

The screenshot shows the ORKG Contribution editor interface. At the top, there is a navigation bar with the ORKG logo, menu items (View, Tools, About), a user profile dropdown (NFDI4DataScience), a search bar, and a '+ Add new' button. A yellow warning banner states: "Warning: You are using a testing environment. Data you enter in the system can be deleted without any notice." Below this is the 'Contribution editor' section with 'View comparison' and '+ Add contribution' buttons. The main area is a table with a 'Properties' column on the left and a 'Value' column on the right. The 'research question' property is selected and its value is being edited. Below the table are buttons for '+ Add property', a lightbulb icon, and '+ Templates'. At the bottom, there is a 'Suggested properties' section with a '+ research problem' button and a 'Problem' button.

Properties	Value
data analysis method	thematic analysis
data collection method	interview
research question	What is the perception and current treatment of NFRs in ML in industry?
threats to validity	

If you have more research questions, **repeat** the process.

Otherwise, continue with the next property.

6. Describe a Paper: Enter the Threats to Validity

The screenshot shows the ORKG Contribution editor interface. A modal window titled "View existing resource: Threats to Validity" is open, displaying the "Applied template: Threats to Validity". The modal contains four rows, each with a plus sign button: "Construct validity", "Internal validity", "External validity", and "Conclusion validity". The background shows the "Contribution editor" with a "Properties" section containing links for "data analysis method", "data collection method", "research question", and "threats to validity".

1. Click the +
2. Enter construct validity
3. Click the +
4. Enter internal validity
5. Click the +
6. Enter external validity
7. Click the +
8. Enter conclusion validity

6. Describe a Paper: Enter the Threats to Validity

The screenshot shows the ORKG Contribution editor interface. A modal dialog titled "View existing resource: Threats to Validity" is open. The dialog has a header with "Open resource" and a close button. Below the header are buttons for "Templates", "Help", and "Preferences". The main content area shows "Applied template: Threats to Validity". There are four rows of property settings, each with a "Boolean" dropdown, a "True" value, a dropdown arrow, a lightbulb icon, and "Cancel" and "Create" buttons. The properties are: Construct validity, Internal validity, External validity, and Conclusion validity. At the bottom of the dialog is a "+ Add property" button with a lightbulb icon. The background shows the "Contribution editor" with a "Properties" list on the left containing "data analysis method", "data collection method", "research question", and "threats to validity". At the bottom, there is a "Suggested properties" section with a "+ research problem" button and a "Problem" button.

1. Click “Create”
2. Click “Create”
3. Click “Create”
4. Click “Create”
5. Close dialog

6. Describe a Paper: Enter the Threats to Validity

The screenshot shows the ORKG Contribution editor interface. At the top, there is a navigation bar with the ORKG logo, menu items (View, Tools, About), a user profile icon, and a search bar. A yellow warning banner states: "Warning: You are using a testing environment. Data you enter in the system can be deleted without any notice." Below this is the "Contribution editor" header with "View comparison" and "Add contribution" buttons. The main content area is a table with a left sidebar for properties and a right column for values. The "threats to validity" property is selected and contains the text: "Construct validity: true, Internal validity: true, External validity: true, Conclusion validity: true". Other properties include "data analysis method" (thematic analysis), "data collection method" (interview), and "research question" (What is the perception and current treatment of NFRs in ML in industry?). At the bottom, there are buttons for "Add property", "Templates", and "Suggested properties" (research problem).

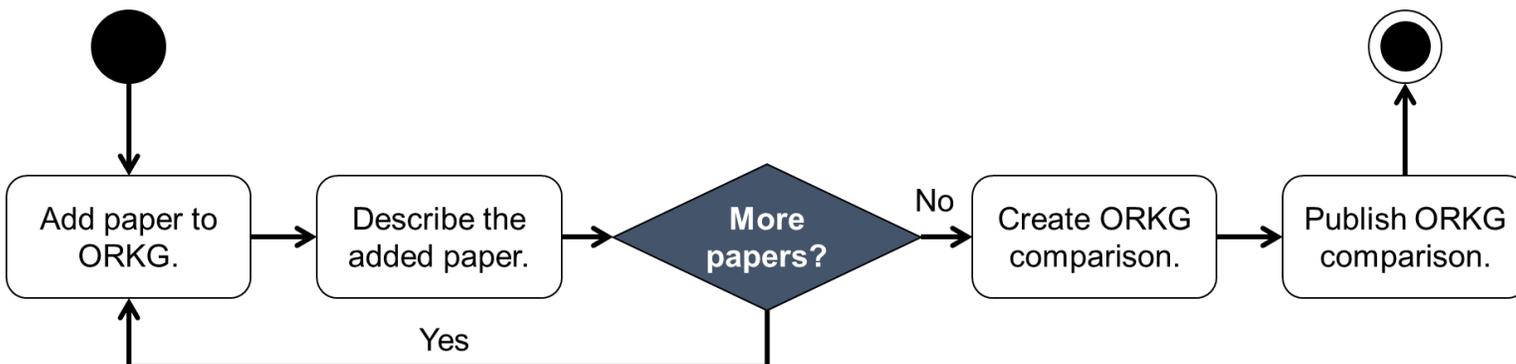
Properties	Values
data analysis method	thematic analysis
data collection method	interview
research question	What is the perception and current treatment of NFRs in ML in industry?
threats to validity	Construct validity: true, Internal validity: true, External validity: true, Conclusion validity: true

Now, we have described our first paper regarding its reported empirical research.

We can repeat the entire process for the next paper or work collaboratively!

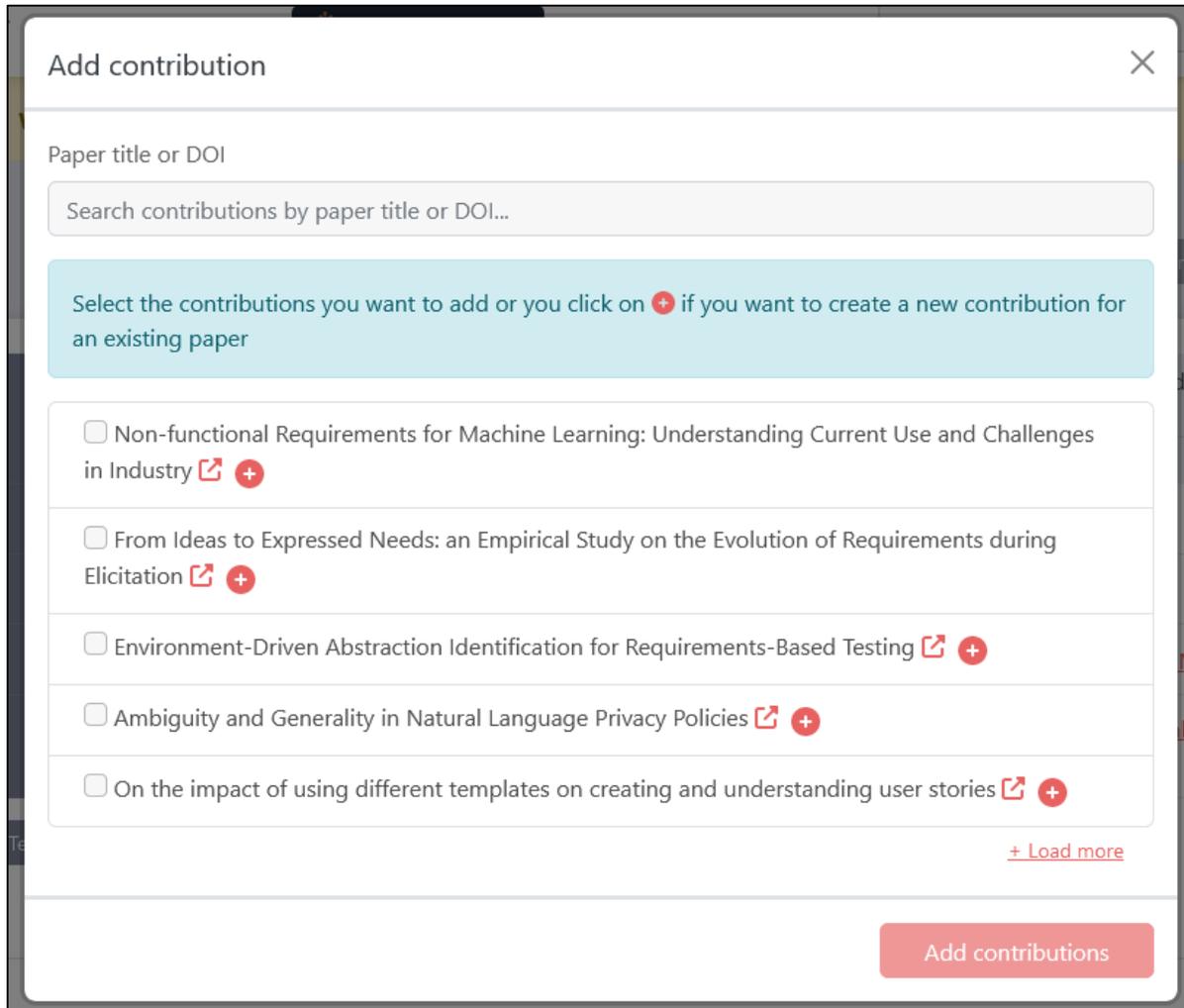
7. Add the Next Paper

1. Click on “Add contribution”



Remark:
Now, we use a paper called **“From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation”** already described with the ORKG template.

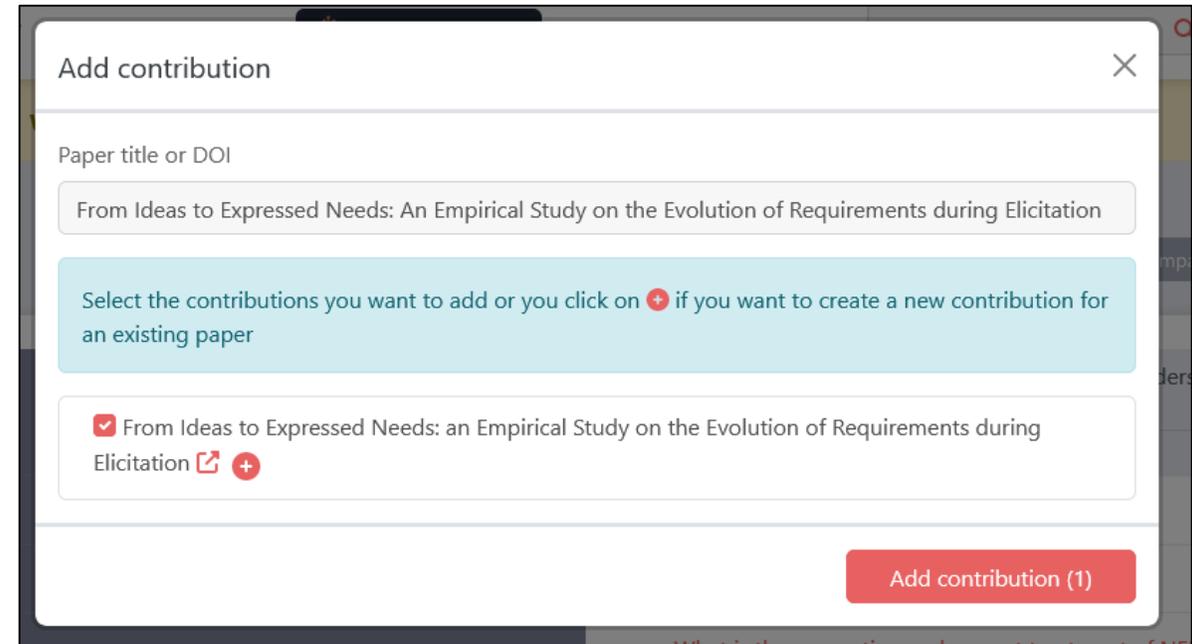
7. Add the Next Paper: Enter Paper Title



The screenshot shows a dialog box titled "Add contribution" with a close button (X) in the top right corner. Below the title is a search bar labeled "Paper title or DOI" with the placeholder text "Search contributions by paper title or DOI...". A light blue instruction box contains the text: "Select the contributions you want to add or you click on + if you want to create a new contribution for an existing paper". Below this is a list of five papers, each with an unchecked checkbox, a red square icon containing a white plus sign, and a red circular icon containing a white plus sign. The papers are:

- Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry
- From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation
- Environment-Driven Abstraction Identification for Requirements-Based Testing
- Ambiguity and Generality in Natural Language Privacy Policies
- On the impact of using different templates on creating and understanding user stories

At the bottom right of the list is a red link "+ Load more". At the bottom center of the dialog box is a red button labeled "Add contributions".



The screenshot shows the same "Add contribution" dialog box, but with the checkbox for the second paper, "From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation", checked. The red square icon and red circular icon are still present. At the bottom right of the dialog box is a red button labeled "Add contribution (1)".

1. Enter paper title
2. Select the checkbox
3. Click "Add contribution (1)"

7. Add the Next Paper: Result

Contribution editor
View comparison
+ Add contribution

	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry ✕	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation ✕
Properties ✓ Saved	Contribution 1	Contribution 1
data analysis method	thematic analysis	thematic analysis
data collection method	interview	experiment
research question	What is the perception and current treatment of NFRs in ML in industry?	What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? How much is the difference in terms of documented requirements and roles with respect to initial ideas?
threats to validity	Construct validity: true, Internal validity: true, External validity: true, Conclusion validity: true	Construct validity: true, Internal validity: true, External validity: true, Conclusion validity: false

Now, we have two described papers.

If we want, we can add further ones or we can create an ORKG comparison.

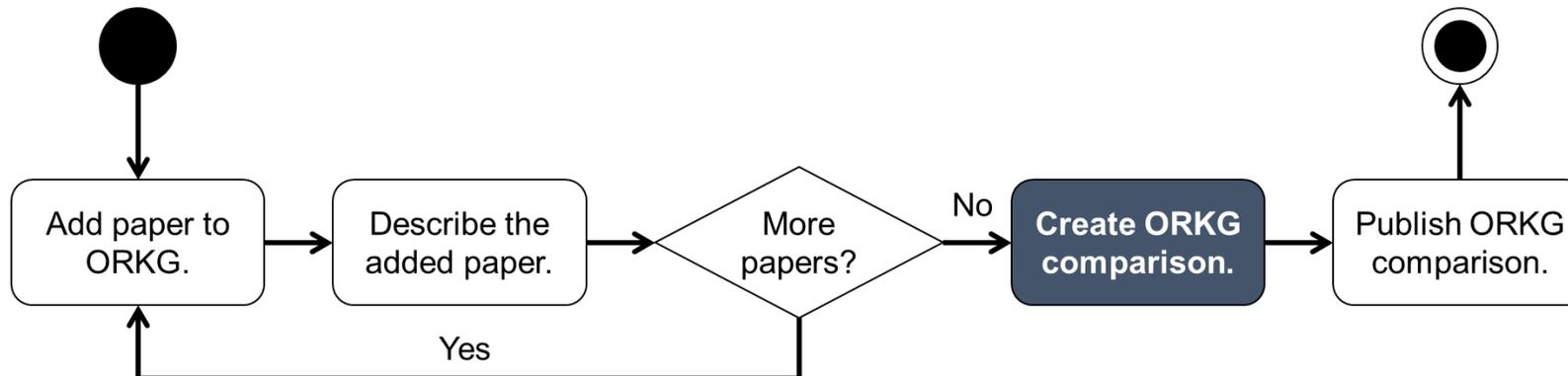
8. Create ORKG Comparison

Contribution editor

View comparison Add contribution

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation
	Contribution 1	Contribution 1
data analysis method	thematic analysis	thematic analysis inferential statistics descriptive statistics
data collection method	interview	experiment
research question	What is the perception and current treatment of NFRs in ML in industry?	What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? How much is the difference in terms of documented requirements and roles with respect to initial ideas?
threats to validity	Construct validity: true, Internal validity: true, External validity: true, Conclusion validity: true	Construct validity: true, Internal validity: true, External validity: true, Conclusion validity: false

1. Click on “View comparison”



8. Create ORKG Comparison: Result and Improvements

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
data.analysis.method	Data Analysis Method	Data Analysis Method
data.analysis.method/data.analysis.method/method*	thematic analysis	descriptive statistics inferential statistics thematic analysis
data.collection.method	Data Collection Method	Data Collection Method
data.collection.method/data.collection.method		
method*	interview	experiment
number.of.participants*	10	30
research.question	Research Question	Research Question
research.question/research.question		
hidden.in.text*	✗	✗
highlighted.in.text*	✓	✓
question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
threats.to.validity	Threats to Validity	Threats to Validity
threats.to.validity/threats.to.validity		
conclusion.validity*	✓	✗
construct.validity*	✓	✓
external.validity*	✓	✓
internal.validity*	✓	✓

Now, we created our ORKG comparison.

Options before publishing the ORKG comparison:

1. Edit the ORKG comparison by ordering the rows
Edit → Drag & Drop the property cells as required
2. Select properties we want to show.
Actions → Select properties → Disable checkboxes of properties to hide
3. Save ORKG comparison as a draft for later
Actions → Save as draft → Enter title → Save → Draft in “My account”

9. Improve ORKG Comparison: Order Rows

Comparison | 2 contributions

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
data.analysis.method	Data Analysis Method	Data Analysis Method
data.analysis.method/data.analysis.method/method*	thematic analysis	descriptive statistics
		inferential statistics
		thematic analysis
data.collection.method	Data Collection Method	Data Collection Method
data.collection.method/data.collection.method		
method*	interview	experiment
number.of.participants*	10	30
research.question	Research Question	Research Question
		Research Question
		Research Question
research.question/research.question		
hidden.in.text*	✗	✗
		✗
		✗
highlighted.in.text*	✓	✓
		✓
		✓
question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas?
		What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
		What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
threats.to.validity	Threats to Validity	Threats to Validity
threats.to.validity/threats.to.validity		
conclusion.validity*	✓	✗
construct.validity*	✓	✓
external.validity*	✓	✓
internal.validity*	✓	✓

ORKG | View | Tools | About | NFDI4DataScience

Search... + Add new

Comparison | 2 contributions

Stop editing | Add contribution | Visualize | Actions

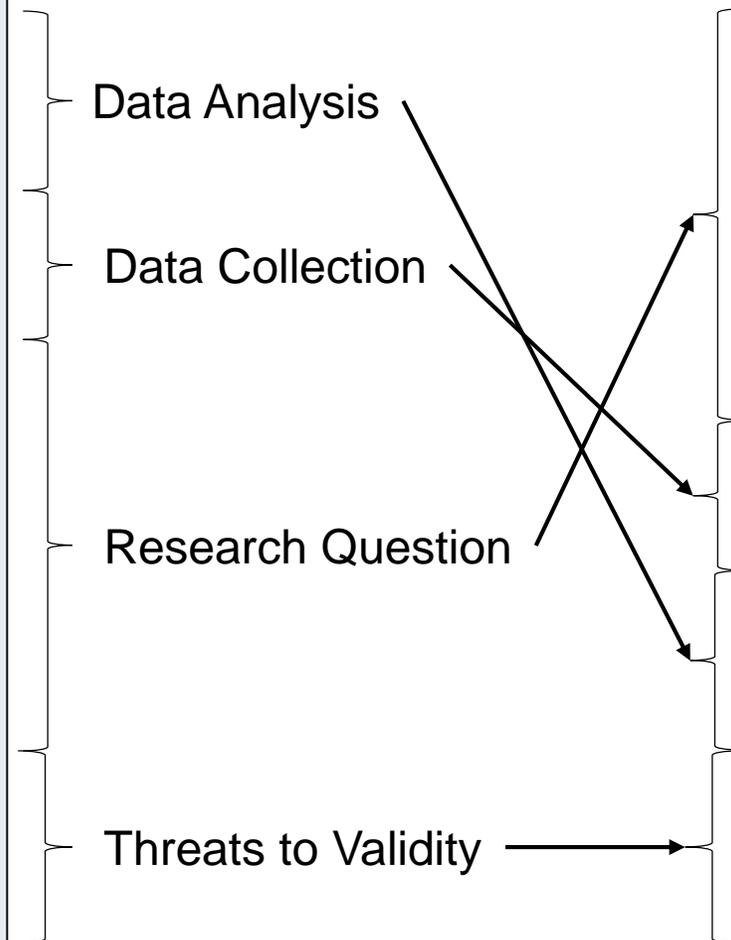
Edit mode

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research question		
data.analysis	Data Analysis Method	Data Analysis Method
		Data Analysis Method
		Data Analysis Method
data.analysis.method/data.analysis.method/method*	thematic analysis	descriptive statistics
		inferential statistics
		thematic analysis

1. Click on “Edit”
2. Click on the grey property cell you want to move
3. Drag & Drop the property cell where you want

9. Improve ORKG Comparison: Order Rows – Result

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
data_analysis_method	Data Analysis Method	Data Analysis Method
data_analysis_method/data_analysis_method/method*	thematic analysis	Data Analysis Method descriptive statistics inferential statistics thematic analysis
data_collection_method	Data Collection Method	Data Collection Method
data_collection_method/data_collection_method		
↳ method*	interview	experiment
↳ number_of_participants*	10	30
research_question	Research Question	Research Question
research_question/research_question		Research Question
↳ hidden_in_text*	✗	✗
↳ highlighted_in_text*	✓	✗
↳ question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
threats_to_validity	Threats to Validity	Threats to Validity
threats_to_validity/threats_to_validity		
↳ conclusion_validity*	✓	✗
↳ construct_validity*	✓	✓
↳ external_validity*	✓	✓
↳ internal_validity*	✓	✓

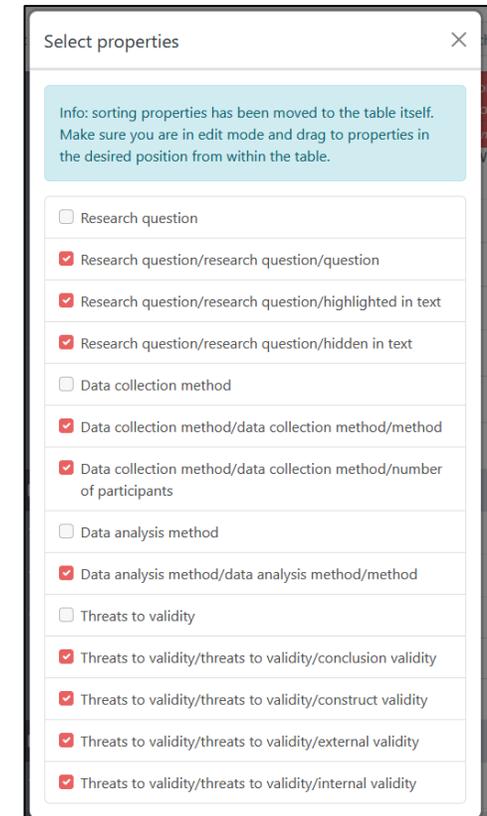
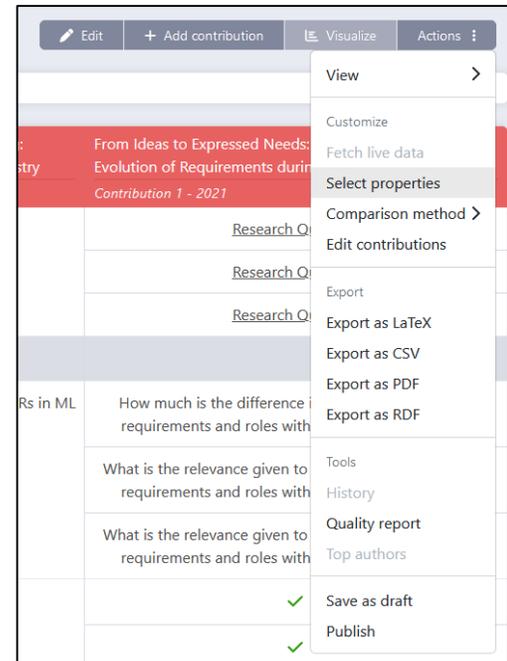


Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research_question	Research Question	Research Question
research_question/research_question		Research Question
↳ question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
↳ highlighted_in_text*	✓	✓
↳ hidden_in_text*	✗	✗
data_collection_method	Data Collection Method	Data Collection Method
data_collection_method/data_collection_method		
↳ method*	interview	experiment
↳ number_of_participants*	10	30
data_analysis_method	Data Analysis Method	Data Analysis Method
data_analysis_method/data_analysis_method/method*	thematic analysis	Data Analysis Method descriptive statistics inferential statistics thematic analysis
threats_to_validity	Threats to Validity	Threats to Validity
threats_to_validity/threats_to_validity		
↳ conclusion_validity*	✓	✗
↳ construct_validity*	✓	✓
↳ external_validity*	✓	✓
↳ internal_validity*	✓	✓

9. Improve ORKG Comparison: Select Properties

Comparison | 2 contributions

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research question	Research Question	Research Question
research question/research question		
question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas?
highlighted in text	✓	✓ ✓ ✓
hidden in text	✗	✗ ✗ ✗
data collection method	Data Collection Method	Data Collection Method
data collection method/data collection method		
method*	interview	experiment
number of participants*	10	30
data analysis method	Data Analysis Method	Data Analysis Method
data analysis method/data analysis method/method*	thematic analysis	descriptive statistics inferential statistics thematic analysis
threats to validity	Threats to Validity	Threats to Validity
threats to validity/threats to validity		
conclusion validity*	✓	✗
construct validity*	✓	✓
external validity*	✓	✓
internal validity*	✓	✓

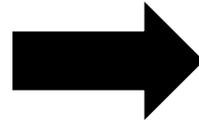


1. Click on “Actions”
2. Click on “Select properties”
3. Disable all checkboxes of properties you want to hide
4. Close dialog

9. Improve ORKG Comparison: Select Properties – Result

Comparison | 2 contributions

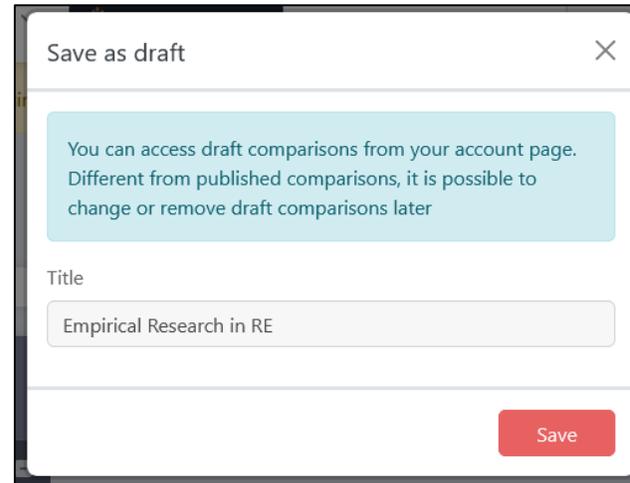
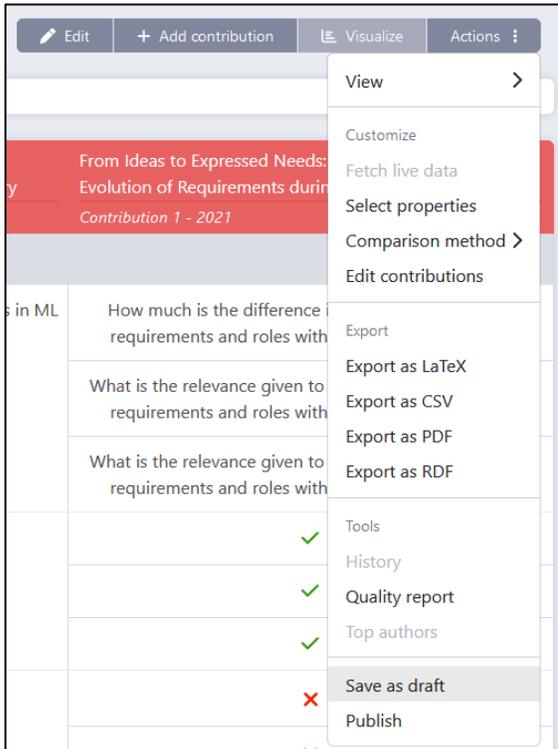
Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research question	Research Question	Research Question
research question/research question		
↳ question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas?
↳ highlighted in text*	✓	✓ ✓ ✓
↳ hidden in text*	✗	✗ ✗ ✗
data collection method	Data Collection Method	Data Collection Method
data collection method/data collection method		
↳ method*	interview	experiment
↳ number of participants*	10	30
data analysis method	Data Analysis Method	Data Analysis Method
data analysis method/data analysis method/method*	thematic analysis	descriptive statistics inferential statistics thematic analysis
threats to validity	Threats to Validity	Threats to Validity
threats to validity/threats to validity		
↳ conclusion validity*	✓	✗
↳ construct validity*	✓	✓
↳ external validity*	✓	✓
↳ internal validity*	✓	✓



Comparison | 2 contributions

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research question/research question		
↳ question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas?
↳ highlighted in text*	✓	✓ ✓ ✓
↳ hidden in text*	✗	✗ ✗ ✗
data collection method/data collection method		
↳ number of participants*	10	30
↳ method*	interview	experiment
data analysis method/data analysis method/method*	thematic analysis	descriptive statistics inferential statistics thematic analysis
threats to validity/threats to validity		
↳ conclusion validity*	✓	✗
↳ construct validity*	✓	✓
↳ external validity*	✓	✓
↳ internal validity*	✓	✓

9. Improve ORKG Comparison: Save as Draft

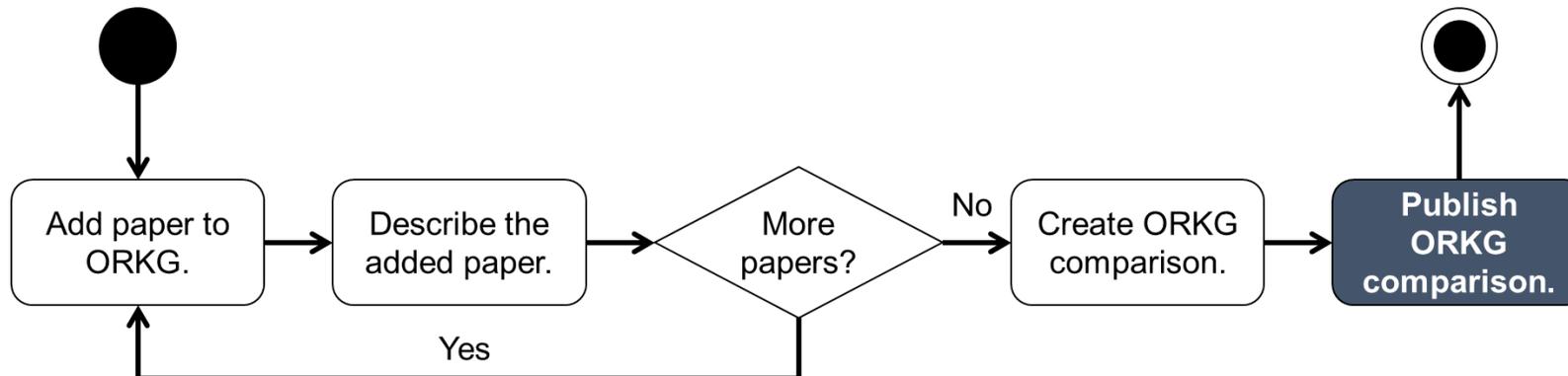
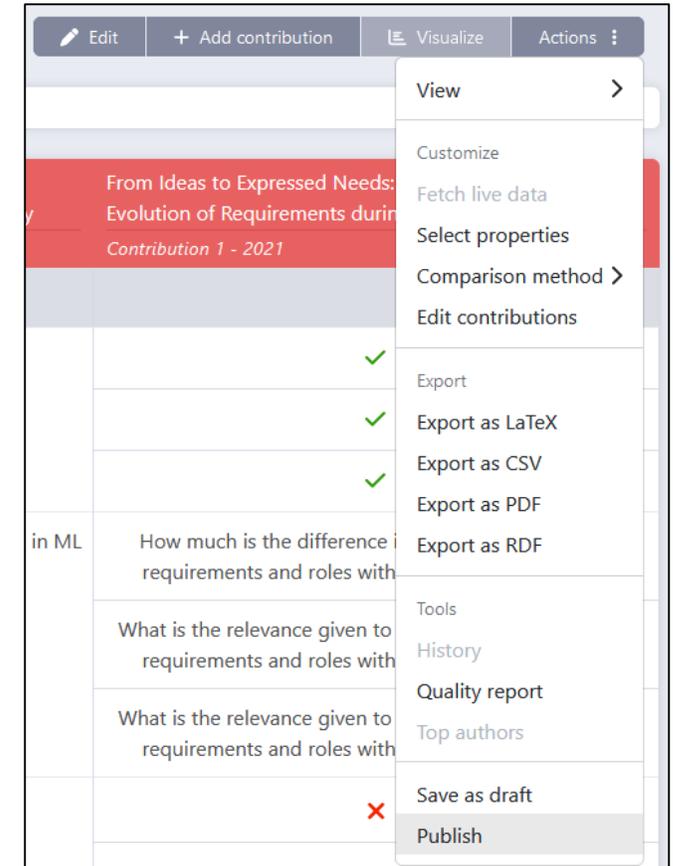


1. Click on “Actions”
2. Click on “Save as draft”
3. Enter a title and click on “Save”
4. The draft is saved in your account

10. Publish ORKG Comparison

Comparison | 2 contributions

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry <i>Contribution 1 - 2021</i>	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation <i>Contribution 1 - 2021</i>
research question/research question		
↳ highlighted in text*	✓	✓
↳ question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas? What is the relevance given to the different categories of requirements and roles with respect to initial Ideas?
↳ hidden in text*	✗	✗



1. Click on “Actions”
2. Click on “Publish”

10. Publish ORKG Comparison

Publish comparison

A published comparison is made public to other users. The state of the comparison is saved and a persistent link is created.

Title ?

Description ?

Research field ?

Creators ?

Assign a DOI to the comparison ?

Reference (optional) ?

Sustainable development goals (optional) ?

Conference (optional) ?

Publish

1. Fill out the dialog
2. Click on “Publish”

Remark:

You can add

- A **DOI** for citation (can also be done later)
- Additional **References**
- Related **Sustainable development goals**
- ORKG partner **Conferences**, such as **REFSQ'25**, that award the **Best ORKG Comparison Award**

Publish comparison

A published comparison is made public to other users. The state of the comparison is saved and a persistent link is created.

Title ?

Description ?

Research field ?

Creators ?

Assign a DOI to the comparison ?

Reference (optional) ?

Sustainable development goals (optional) ?

Conference (optional) ?

a conference
asdadfa
31st International Working Conference on Requirement Engineering: Foundation for Software Quality (REFSQ'25)

10. Publish ORKG Comparison: Result

Comparison | 2 contributions

An Overview of Empirical Research in Requirements Engineering ☆

May 2024 • Oliver Karras

This comparison shows an overview of empirical research reported in publication of the IEEE International Requirements Engineering Conference 2021 regarding the topics research question, data collection, data analysis, and threats to validity.

Properties	Non-functional Requirements for Machine Learning: Understanding Current Use and Challenges in Industry Contribution 1 - 2021	From Ideas to Expressed Needs: an Empirical Study on the Evolution of Requirements during Elicitation Contribution 1 - 2021
research question/research question		
highlighted in text*	✓	✓
question*	What is the perception and current treatment of NFRs in ML in industry?	How much is the difference in terms of documented requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas? What is the relevance given to the different categories of requirements and roles with respect to initial ideas?
hidden in text*	✗	✗
data collection method/data collection method		
method*	interview	experiment
number of participants*	10	30
data analysis method/data analysis method/method*	thematic analysis	descriptive statistics inferential statistics thematic analysis
threats to validity/threats to validity		
conclusion validity*	✓	✗
construct validity*	✓	✓
external validity*	✓	✓
internal validity*	✓	✓

Added by: Oliver Karras

Assign to observatory

Now, we **published our ORKG comparison**, a stable version that can be maintained, extended, updated, and published as new versions.

Options after publishing the ORKG comparison:

1. Add visualizations
2. Add DOI later and export citation
3. Use the quality report to get feedback from other researchers
4. Fetch live data for a new draft or published version
5. Fetch data for later analysis with different interfaces

Remark:

These options are only demonstrated live in the tutorial. If you need help, do not hesitate to contact the ORKG team, especially **Oliver Karras** (oliver.karras@tib.eu).

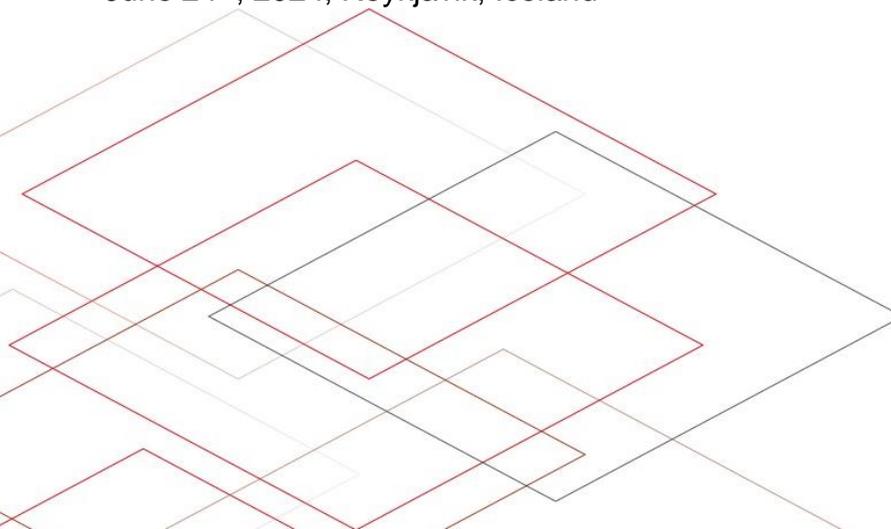
Summary, Reflection, and Closing

Oliver Karras, Alessio Ferrari, Davide Fucci, and Davide Dell'Anna

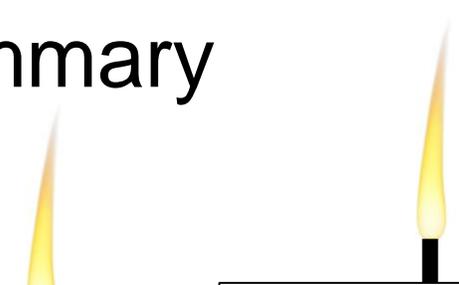
oliver.karras@tib.eu, alessio.ferrari@isti.cnr.it, davide.fucci@bth.se, d.dellanna@uu.nl

32nd IEEE International Requirements Engineering 2024 Conference – Exploring New Horizons: Expanding the Frontiers of Requirements Engineering

June 24th, 2024, Reykjavik, Iceland



Summary



Mining User Requirements from Application Store Reviews Using Frame Semantics

Nishaat Jha and Anas Mahmood^[1]

The Division of Computer Science and Engineering, Louisiana State University, Baton Rouge, LA 70803, USA
njha@lsu.edu, mahmood@ce.lsu.edu

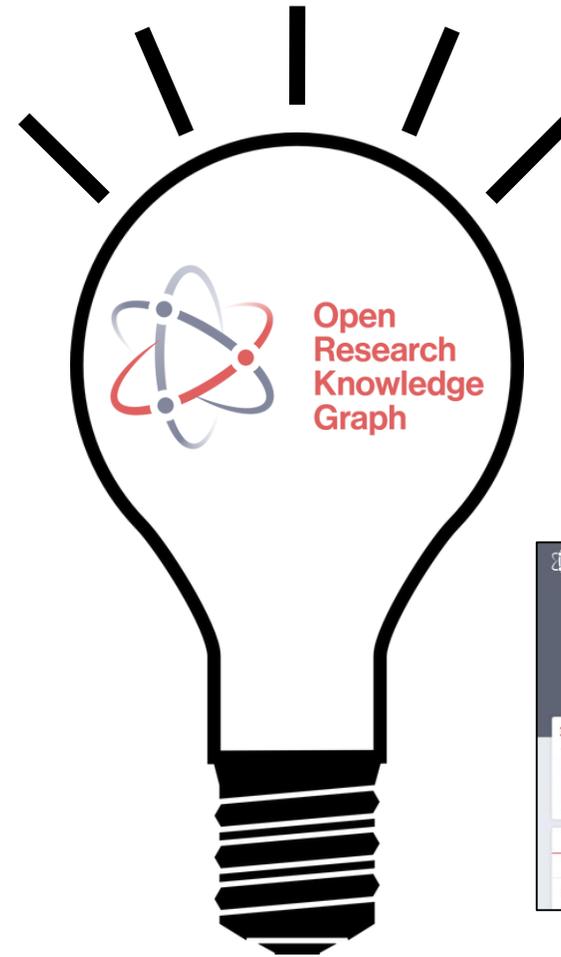
Abstract. Context and motivation: Research on mining user reviews in mobile application (app) stores has noticeably advanced in the past few years. The majority of the proposed techniques rely on classifying the textual description of user reviews into different categories of technically informative user requirements and uninformative feedback. **Question/Problem:** Relying on the textual attributes of reviews often produces high dimensional models. This increases the complexity of the classifier and can lead to overfitting problems. **Principal idea/solution:** We propose a novel semantic approach for app review classification. The proposed approach is based on the notion of semantic role labeling, or characterizing the lexical meaning of text in terms of semantic frames. Semantic frames help to generalize from text (individual words) to more abstract scenarios (contexts). This reduces the dimensionality of the data and enhances the predictive capabilities of the classifier. Three datasets of user reviews are used to conduct our experimental analysis. Results show that semantic frames can be used to generate lower dimensional and more accurate models in comparison to text classification methods. **Contribution:** A novel semantic approach for extracting user requirements from app reviews. The proposed approach enables a more efficient classification process and reduces the chance of overfitting.

Keywords: Requirements elicitation · Application stores · Classification

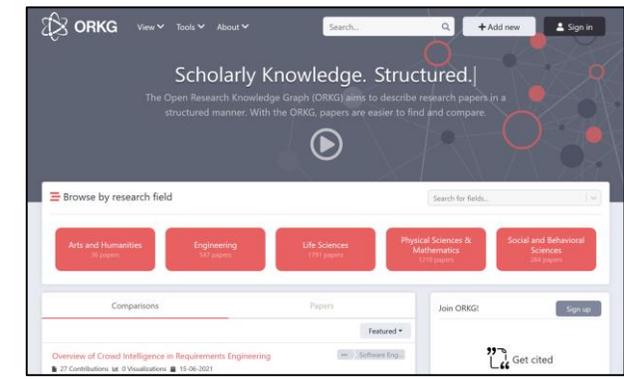
1 Introduction

Mobile application markets, or app stores (e.g., Google Play and Apple App Store), represent a unique model of service-oriented business. Such platforms have created an unprecedented opportunity for app developers to directly monitor the opinions of a large population of end-users of their software [2]. Through app stores feedback services, app users can directly share their experience in the form of textual reviews and meta-data (e.g., star ratings). Analyzing large datasets of app store reviews has revealed that they contain substantial amounts of up-to-date technical information. Such information can be leveraged by app developers to help them maintain and sustain their apps in a highly-competitive

© Springer International Publishing AG 2017
© Gölubaslar and A. Pinar (Eds.) REEPQ 2017, LNCS 10153, pp. 273–287, 2017.
DOI: 10.1007/978-3-319-54064-2_20



<https://orkg.org>



Let's bring scholarly communication and open science in Requirements Engineering to the 21st century!

Recap of the Tutorial

Session	Time	Table of Content	Style	Speaker
Theoretical	09:00 - 09:25	1. Welcome (5 min) 2. Introduction to open science in RE (10 min) 3. Introduction to the ORKG (10 min)	Presentation Presentation Presentation	All organizers Alessio Ferrari Oliver Karras
Practical	09:25 - 10:15	4. Create a FAIR-annotated publication for the ORKG (50 min) 4.1 Set up an Overleaf project for an exemplary publication 4.2 Use the LaTeX package SciKGT _E X to annotate the publication 4.3 Generate PDF with embedded FAIR scientific information 4.4 Optional: Upload the FAIR-annotated publication to the ORKG	Exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise	Oliver Karras All organizers All organizers All organizers All organizers
Break	10:15 - 10:45	Coffee break		
Practical	10:45 - 11:45	5. Use the ORKG based on a RE use case (60 min) 5.1 Add an exemplary publication to the ORKG 5.2 Describe the scientific information of the publication in the ORKG 5.3 Create an ORKG comparison of the publications added by participants 5.4 Publish the created ORKG comparison as a citable digital artifact 5.5 Optional: Create visualizations for the created ORKG comparison 5.6 Optional: Retrieve the information with the SPARQL endpoint	Exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise Sub-exercise	Oliver Karras All organizers All organizers All organizers All organizers All organizers All organizers
Feedback	11:45 - 12:15	6. Reflection of the tutorial with the participants (25 min) 7. Farewell and closing (5 min)	Discussion Presentation	All organizers All organizers

Reflection and Feedback

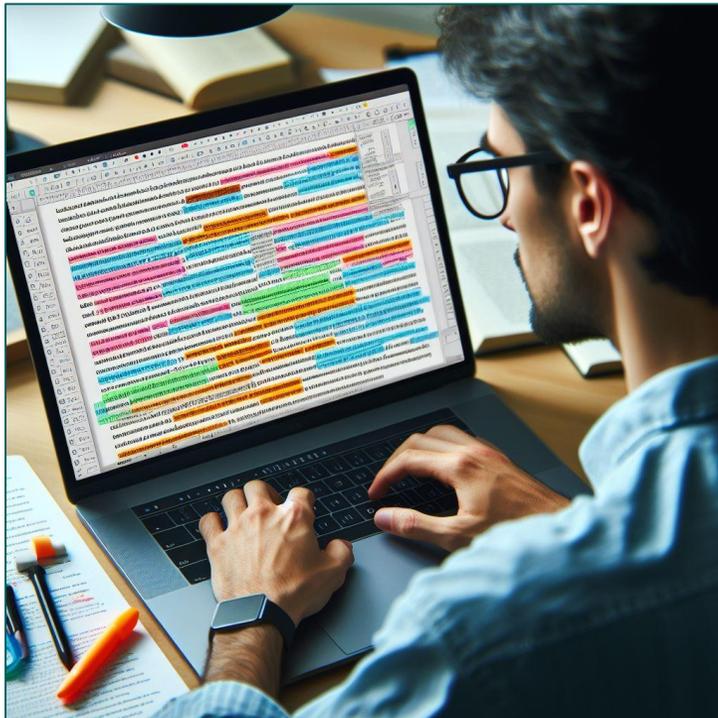
Your thoughts, please!

Teaser: REFSQ'25 Open Science Competition

Challenge 1:

Annotate your REFSQ'25 paper with SciKGTEx.

The accepted paper, best annotated with SciKGTEx, will be awarded the **Best ORKG Annotation Award** (prize: 100€).



SciKGTEx

Challenge 2:

Enrich your RESFQ'25 paper with an ORKG comparison.

The accepted paper, enriched with the best ORKG comparison, will be awarded the **Best ORKG Comparison Award** (prize: 200€).



Open Research Knowledge Graph

Divide and Conquer the EmpiRE:
A Community-Maintainable Knowledge Graph of Empirical Research in Requirements Engineering

Oliver Karas*, Felix Wernlein[†], Jill Klünder[†] and Sören Auer*[†]
*TIB - Leibniz Information Centre for Science and Technology, Hannover, Germany
Email: {oliver.karas, soeren.auer}@tib.eu
[†]Leibniz University Hannover, Hannover, Germany
Email: felix.wernlein@stud.uni-hannover.de, jill.kluender@inf.uni-hannover.de, auer@13s.de

A Comparison of Scientific Publications on the State of Empirical Research in Requirements Engineering and Software Engineering ★

November 2023 | Oliver Karas | Felix Wernlein | Jill Ann-Christin Klünder | Sören Auer

This comparison provides an overview of scientific publications that have investigated primary studies in requirements engineering and software engineering to give a snapshot of the "current" state of empirical research in requirements engineering and software engineering. In particular, the comparison shows for each publication (1) which research fields and topics were investigated, (2) whether and where the extracted and analyzed data is available, and (3) which method was used to determine the state, including further details about the respective method.

DOI: <https://doi.org/10.48366/R650023>

Properties	Empirical research in requirements engineering: trends and opportunities - 2016	Empirical research methodologies and studies in Requirements Engineering: How far did we come? - 2014	A Survey on Empirical Requirements Engineering Research Practices - 2012	Evidence-Based Structuring and Evaluation of Empirical Research in Requirements Engineering: Fundamentals, Framework, Research Map - 2010	An Anal. Requires Data - Empirical
research problem	empirical research in requirements engineering	empirical research in requirements engineering	empirical research in requirements engineering	empirical research in requirements engineering	empirical
research field investigated	Requirements Engineering	Requirements Engineering	Requirements Engineering	Requirements Engineering	Req.
topic investigated	bibliographic metadata context data collection	bibliographic metadata research topic theory	context data analysis data collection	context research method result	

Abstract—(Background.) Empirical research in requirements engineering (RE) is a constantly evolving topic, with a growing (systematic) literature reviews, and even examine overlapping periods, venues, and themes (cf. Table 1) [16], they have not collaborated to build on and update earlier works, which are known challenges of literature reviews [17]–[20]. Overcoming these challenges is critical to ensure the quality, reliability, and timeliness of research results from literature reviews [19], [21]. Recent research addresses these challenges by focusing on when and how to update (systematic) literature reviews in SE and its subfields [4], [21]–[23]. While these works mainly provide social and economic decision support and guidance for updating literature reviews [4], [20], the underlying problem is the unavailability of the extracted and analyzed data, corresponding to open science in SE [23], [24]. Unavailable data complicates collaboration among researchers and updating literature reviews, as the entire data collection, extraction, and analysis must be repeated and expanded for comprehensive results. Researchers need support in the form of technical infrastructures and services to conduct sustainable literature reviews so that all data is openly available in the long term [5], [17], [18], [29] according to the Findable, Accessible, Interoperable, and Reusable (FAIR) data principles [25], [26]. For this purpose, the data must be organized in a flexible, fine-grained, context-sensitive, and semantic representation to be understandable, processable, and usable by humans and machines [5], [13], [27]. Over the last decade, Knowledge Graphs (KGs) have become an emerging technology in industry and academia as they enable this versatile data representation [28]–[30]. Besides well-known KGs for encyclopedic and factual data, such as *Dispedia* [31] and *WikiData* [32], using so-called Research Knowledge Graphs (RKGs) for scientific data is a rather new approach [28], [29], [33]. RKGs include bibliographic metadata, e.g., titles, authors, and venues, as well as scientific data, e.g., research designs, methods, and results [34]–[39]. They are a promising technology to sustainably organize scientific data so that the data is openly available for long-term collaborations [27], [40]. We examine the use of RKGs as technical infrastructure by building, publishing, and evaluating an initial KG of Empirical research in RE (KG-EmpIRE). Similar to Fratini et al. [41], our long-term goal is to continuously maintain, your personal use. Not for redistribution. The definitive version or record was published in the proceedings of 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), <https://doi.org/10.1109/ESEM56168.2023.10304795>.

Looking to the Future: ORKG Ask

The screenshot shows the ORKG Ask website interface. At the top left is the ORKG Ask logo. Navigation links for 'Search' and 'My library' are visible. A 'Sign in' button is in the top right. The main heading reads 'Find research you are *actually* looking for'. Below this is a search bar with the placeholder text 'Ask your question...'. A descriptive box states: 'ORKG Ask is a scholarly search and exploration system powered by **Vector Search, Large Language Models** and **Knowledge Graphs**.' To the right, a large number '76.430.670' is displayed above the word 'Papers'. A 'Getting started' section lists five research questions in red text: 'What are the long-term effects of income inequality on community well-being?', 'What role does trust play in building resilient communities?', 'What are the ethical implications of using artificial intelligence in decision-making processes?', 'What is the significance of higher-dimensional algebra?', and 'What are the social implications of widespread adoption of autonomous vehicles?'. Below this is a logo for 'EULIST' with the text 'ORKG Ask is brought to you by'. At the bottom, a footer contains links for 'Contact', 'Data Protection', 'Accessibility', 'Imprint', 'Changelog', and 'Version: 1.6.0'.

The ORKG Ask is a scientific **search and exploration system** that helps researchers **find** the **research articles** they are really looking for.

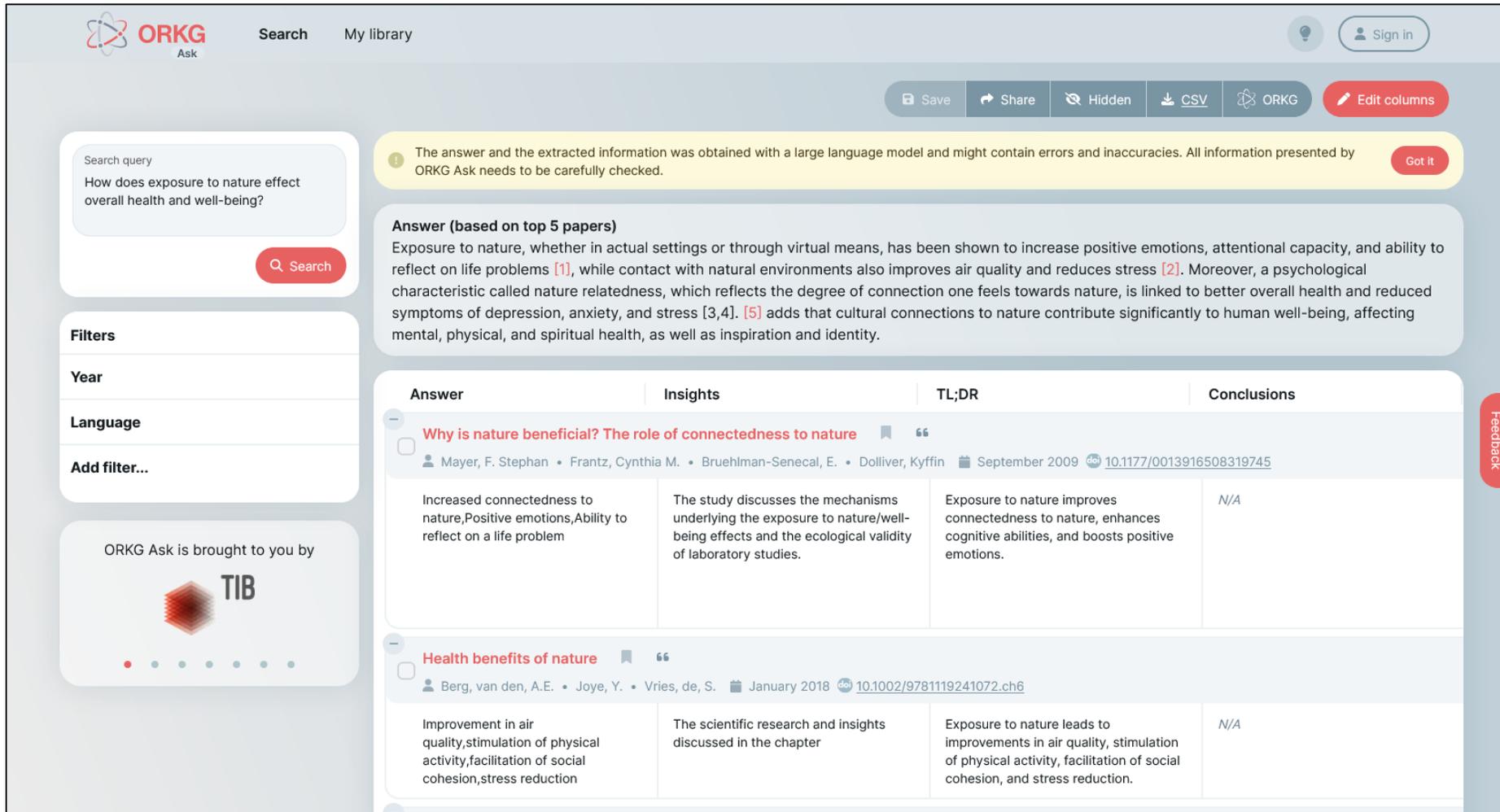
Ask a **question**.

Get an **answer** with **references**.

<https://ask.orkg.org/>

(Released: May16th, 2024)

Looking to the Future: ORKG Ask



Search query: How does exposure to nature effect overall health and well-being?

Filters: Year, Language, Add filter...

Warning: The answer and the extracted information was obtained with a large language model and might contain errors and inaccuracies. All information presented by ORKG Ask needs to be carefully checked.

Answer (based on top 5 papers): Exposure to nature, whether in actual settings or through virtual means, has been shown to increase positive emotions, attentional capacity, and ability to reflect on life problems [1], while contact with natural environments also improves air quality and reduces stress [2]. Moreover, a psychological characteristic called nature relatedness, which reflects the degree of connection one feels towards nature, is linked to better overall health and reduced symptoms of depression, anxiety, and stress [3,4]. [5] adds that cultural connections to nature contribute significantly to human well-being, affecting mental, physical, and spiritual health, as well as inspiration and identity.

Answer	Insights	TL;DR	Conclusions
<p>Why is nature beneficial? The role of connectedness to nature</p> <p>Mayer, F. Stephan • Frantz, Cynthia M. • Bruehlman-Senecal, E. • Dolliver, Kyffin • September 2009 • 10.1177/0013916508319745</p> <p>Increased connectedness to nature, Positive emotions, Ability to reflect on a life problem</p>	<p>The study discusses the mechanisms underlying the exposure to nature/well-being effects and the ecological validity of laboratory studies.</p>	<p>Exposure to nature improves connectedness to nature, enhances cognitive abilities, and boosts positive emotions.</p>	N/A
<p>Health benefits of nature</p> <p>Berg, van den, A.E. • Joye, Y. • Vries, de, S. • January 2018 • 10.1002/9781119241072.ch6</p> <p>Improvement in air quality, stimulation of physical activity, facilitation of social cohesion, stress reduction</p>	<p>The scientific research and insights discussed in the chapter</p>	<p>Exposure to nature leads to improvements in air quality, stimulation of physical activity, facilitation of social cohesion, and stress reduction.</p>	N/A

1. Enter a natural language **question** in the **UI**
2. **Semantic search** finds the most relevant **papers**
3. **LLMs** extract the required information and create the **answer**
4. **KGs** are used to **support** content extraction, synthesis, and enhancement